Homeobox-containing genes in the nemertean *Lineus*: Key players in the antero-posterior body patterning and in the specification of the visual structures

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel und der Universität Reims
von

Véronique Charpignon
aus
Frankreich

Basel, 2007

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von:

Prof. Dr. Walter J. Gehring

Prof. Dr. Heinrich Reichert

Dr. Michel Tarpin

Basel, den 21. 11. 2006

Prof. Dr. Hans-Peter Hauri
Dekan
Philosophisch-Naturwissenschaftliche Fakultät
Universität Basel

International Cotutelle:
U.F.R. Sciences Exactes et Naturelles Université de Reims Champagne-Ardenne / Philosophisch-Naturwissenschaftliche Fakultät der Universität Basel

Inaugural Dissertation / Thèse de Doctorat
Cellular and Molecular Biology
Submitted in fulfillment of the requirements for the degree of Doctor of Philosophy in life sciences from the University of Basel, Switzerland
and from the University of Reims, France

Véronique CHARPIGNON

Public defense, University of Reims:
November 23rd 2006

Homeobox-containing genes in the nemertean *Lineus*: Key players in the antero-posterior body patterning and in the specification of the visual structures

Under the co-supervison of:
Prof. Walter GEHRING / Dr. Michel TARPIN

Thesis committee:

M. Jacques BIERNE	Professor, University of Reims
M. Walter GEHRING	Professor, University of Basel
Mme Marie KMITA	Doctor, University of Montreal
M. Heinrich REICHERT	Professor, University of Basel
M. Michel TARPIN	Doctor, University of Reims

Short summary

In this thesis, the lophotrochozoan nemertean *Lineus*, which exhibits impressive developmental plasticity and regeneration capacities, was used as a model system in an attempt to reveal to which extent the anteroposterior patterning mechanism and "the eye specification network" are conserved throughout the Bilateria.

The data obtained from the expression patterns of *orthodenticle*-like (*Ls-Otx*) and *caudal*-like (*Ls-Cdx*) genes in *Lineus* are in good agreement with their proposed evolutionarily conserved functions in the specification of the anterior body regions and in the specification of the posterior ones, respectively. We have also shown that *Ls-Cdx* is expressed during posterior regeneration only, whereas *Ls-Otx* is expressed during both, anterior and posterior early regeneration and becomes restricted to the anterior regenerating blastema only one week after the onset of regeneration. Based on its specific expression at the level of the CNS in early regenerating stages, we have proposed that *Otx* could be part of a signaling network responsible for the onset of regeneration in nemertean.

We have previously shown that *LsPax-6* is expressed in developing and regenerating *Lineus* eyes and that it is required for their maintenance in adult *Lineus*. Now, we have obtained data from the three *Ls-Six* genes (*Ls-Six1/2*, *Ls-Six3/6* and *Ls-Six4/5*) that argue for a general involvement of the *Ls-Six* genes in the development and the regeneration of the *Lineus* sensory organs, including the eyes. Hence, evolutionarily conserved key members of the "eye specification network" seems be involved in the specification of the *Lineus* eyes, supporting the hypothesis of a monophyletic origin of the eyes.

1700 characters

5 key words: nemertean / regeneration / evolution / body axis patterning / eye specification

Résumé court

Afin d'estimer le degré de conservation des mécanismes de la mise en place de l'axe antéroposterieur et de la formation des yeux au sein des Bilatériens, nous avons étudié la némerte *Lineus*, un Lophotrochozoaire aux étonnantes capacités de régénération et de régulation morphogénétique.

Les données sur l'expression de l'homologue d'*Orthodenticle* (*Ls-Otx*) et de l'homologue de *caudal (Ls-Cdx)* chez *Lineus* accréditent l'hypothèse d'une conservation au cours de l'évolution de leurs rôles respectifs dans l'acquisition de l'identité céphalique et de l'identité caudale par les extrémités du corps des Bilatériens. Alors que seuls les blastèmes postérieurs expriment *Ls-Cdx*, une expression de *Ls-Otx* est indifféremment retrouvée au niveau des blastèmes antérieurs et postérieurs pendant la première semaine de régénération, avant de n'être limitée qu'aux blastèmes antérieurs. Se basant notamment sur l'expression de *Ls-Otx* par le SNC au cours de la régénération, nous avons proposé qu'*Otx* fasse partie d'une voie de signalisation controlant l'initiation de la régénération chez les némertes.

Une expression de *LsPax-6* est détectée dans les yeux de *Lineus* en développement ou en régénération et nous avons déjà montré qu'elle est nécessaire au maintien des yeux chez l'adute. Les nouveaux résultats obtenus sur les trois gènes *Ls-Six* (*Ls-Six1/2*, *Ls-Six3/6* and *Ls-Six4/5*) suggèrent leurs implications dans le développement et la régénération de nombreux organes sensoriels, et notamment des yeux. Des facteurs clés, à l'origine de la cascade génétique conduisant à l'édification d'un oeil, semblent donc être également conservés chez *Lineus*, renforcant l'hypothèse d'une origine monophylétique de l'oeil.

1698 charactères

5 mots clés: némertes / régénération / evolution / axe antéropostérieur / formation de l'oeil

Abstract

One of the most important breakthroughs in the field of developmental biology has been the discovery of the homeobox and of its widespread phylogenetic conservation. Many homeobox-containing genes encode transcription factors that regulate gene expression during important developmental processes, such as patterning and cell differentiation. Not only their sequences, but often also their expression patterns and their functions are conserved throughout bilaterian animals. Despite specific knowledge from selected model organisms, which belong to the Deuterostomia and the Ecdysozoa, an unified view about the evolutionary conserved developmental mechanisms requires more investigations from the Lophotrochozoa, the third clade of Bilateria, which has been neglected, so far.

We have worked with nermerteans, also called ribbonworms, which are members of the Lophotrochozoa. Because of its evolutionary position, its relative simplicity and impressive developmental plasticity, *Lineus sanguineus*, a marine ribbonworm from the class Anopla, is an attractive system to investigate the specification of the body plan and the mechanism by which differentiated cells maintain or reprogram their identity in a context-dependent manner. In this thesis, *Lineus* was used as a model system in an attempt to reveal to which extent the rostral/ caudal specification of the antero-posterior axis and the eye specification network are conserved throughout the Bilateria.

Although the *Hox* genes play important roles in the antero-posterior specification of the bilaterian body, the most rostral and the most caudal regions of the embryo are specified by *orthodenticle*-like (*Otx*) and *caudal*-like (*Cdx*), respectively. To test whether this is also the case in Lophotrochozoa, we first have characterized the full-length *Ls-Otx* and *Ls-Cdx* genes. Then, we have shown that expression patterns in developing and adult *Lineus* suggest an involvement of *Otx* in the development, the specification and the maintenance of the anterior sensory structures and anterior brain regions. This is in good agreement with the proposed conserved functions of *Otx* among Bilateria. Similarly, the restriction of *Ls-Cdx* expression at the posterior extremity of the developing *Lineus* larva suggest that the presumed conserved role of *Cdx* in the specification of the posterior end of bilaterian embryos could be conserved in *Lineus*. Additionally, we have studied both, the expression patterns and the variation of expression levels of *Ls-Otx* and *Ls-Cdx*

during regeneration. This has revealed that *Ls-Cdx* is specifically up-regulated, during posterior regeneration, only, whereas *Ls-Otx* is up-regulated during both, anterior and posterior regeneration. The *Ls-Otx* expression becomes restricted to the anterior regenerating blastema only one week after the onset of regeneration. As it has been suggested that the CNS plays a crucial role in nemertean regeneration and as *Ls-Otx* is specifically expressed at the tip of the sectioned nerve cord of the early regenerating stages, we propose that *Ls-Otx* could be part of a signaling network responsible for the onset of regeneration. Additional information has been obtained from *Lineus lacteus*, a close relative of *Lineus sanguineus*, which does not exhibit the same regeneration capacities. In the light of the expression pattern of *Otx* in amputated *Lineus lacteus*, we propose that the differences in regeneration capacities between nemertean species could rely on the differences in the capacity of their differentiated cells to de-differentiate in response to signals emitted from *Otx* expressing cells of the nerve cord, rather than in the capacity to emit the signals leading to the onset of regeneration.

In a second project, we have investigated the specification of the visual structures in *L.sanguineus*. Studies in *Drosophila* and vertebrates have revealed that a combinatorial expression of members of the evolutionary conserved "eye specification network" specify the eye field. The key members of this eye specification network are the *Pax-6*, *Six*, *Eyes absent* and *Dachshund* genes. We wanted to know whether this network is involved in the development, maintenance and regeneration of the *Lineus* eyes. At the beginning of this PhD work, it was already known that *LsPax-6* is expressed in the developing eye field in *Lineus*. In addition, we had reported that its inactivation by RNA-mediated gene interference (RNAi) in an adult *L.sanguineus* leads to the disappearance of the adult eyes. To further investigate the specification of the *Lineus* eyes, we have characterized three *Six* genes, *LsSix1/2*, *LsSix3/6* and *LsSix4/5*. Their expressions, especially the one of *LsSix1/2*, suggest an involvement in the development and the regeneration of the *Lineus* eyes. In addition, we have observed a cross-reaction of a *Drosophila* antibody anti-dachshund with the developing *Lineus* eyes. Taken together, these data support the idea that the "eye specification network" could be conserved in nemerteans. This molecular unity underlying eye specification in all bilaterian clades strongly supports the hypothesis of a monophyletic origin of the eyes.

Résumé

Une des découvertes les plus importantes dans le domaine de la biologie du développement a été la mise en évidence de la remarquable conservation phylogénétique de l'homéoboite. Une grande majorité des gènes à homéoboite code pour des facteurs de transcription qui régulent l'expression d'autres gènes au cours de processus développementaux essentiels, comme la détermination des axes et la différentiation cellulaire. Non seulement les séquences des gènes à homéoboite mais également très souvent leurs patrons d'expression ainsi que leurs fonctions sont fortement conservés au sein des Bilatériens. Bien que disposant de nombreuses informations obtenues à partir de quelques organismes modèles, appartenant à seulement deux des trois grands groupes de Bilatériens, les Deutérostomiens et les Ecdysozoaires, de plus amples recherches sur des Lophotrochozoaires, des organismes appartenant au troisième groupe des Bilatériens, sont nécessaires à l'obtention d'une représentation unifiée des mécanismes developpementaux qui ont été conservés au cours de l'évolution au sein des Bilatériens. Afin d'estimer le degré de conservation des mécanismes de la mise en place de l'axe antéroposterieur et de la formation des yeux au sein des Bilatériens, nous avons choisi comme modèle d'étude la némerte *Lineus*. Ce Lophotrochozoaire présente d'étonnantes capacités de régénération et de régulation morphogénétique chez l'adulte, ce qui en fait un modèle de choix pour l'étude des mécanismes de maintien ou de reprogrammation de l'identité de cellules différenciées selon le contexte environnemental.

Bien que les gènes *Hox* soient réputés pour leur implication dans la réalisation du plan d'organisation corporelle le long de l'axe antéro-postérieur, ce sont d'autres gènes à homéoboite, les gènes homologues à *Orthodenticle* (*otd/ Otx*) et à *Caudal* (*cad*/ *Cdx*), qui spécifient respectivement les extrémités antérieures et les extrémitiés postérieures de la majorité des Bilatériens. Afin de s'assurer que tel est également le cas chez les Lophotrochozoaires, nous avons tout d'abord cherché à cloner les gènes *Ls-Otx* et *Ls-Cdx* afin de pouvoir étudier leurs patrons d'expressions. L'étude de ceux-ci a permis de laisser présager du rôle de *Ls-Otx* dans le développement et la régénération des structures sensorielles antérieures et des régions antérieures du cerveau ainsi que du rôle de *Ls-Cdx* dans la spécification de l'extrémité postérieure du corps chez *Lineus*. Ces résultats accréditent l'hypothèse d'une conservation au cours de l'évolution du rôle respectif de ces gènes dans l'acquisition de l'identité céphalique et de l'identité caudale par les extrémités du corps des organismes bilatériens en

développement. De plus, nous avons étudié en détail les patrons d'expressions de ces gènes et la variation de leur niveau d'expression au cours de la régénération. Ceci a permis de montrer que seuls les blastèmes postérieurs expriment *Ls-Cdx*, alors qu'une expression de *Ls-Otx* est indifféremment retrouvée au niveau des blastèmes antérieurs et postérieurs au cours de la première semaine de régénération, avant de n'être limitée qu'aux blastèmes antérieurs. Se basant notamment sur l'expression spécifique de *Ls-Otx* par les cordons nerveux au niveau du plan d'amputation lors de la régénération, nous avons proposé qu'*Otx* fasse partie d'une voie de signalisation contrôlant l'initiation de la régénération chez les némertes.

Dans un second projet, nous avons entrepris l'étude de la spécification des structures visuelles chez *L.sanguineus*. De nombreuses études réalisées chez la Drosophile et chez les vertébrés ont mis en évidence l'existence d'un "réseau génétique de détermination rétinienne", conservé au cours de l'évolution. En effet, l'expression combinée de gènes, tels que *Pax-6*, *Six*, *Eyes absent* et *Dachshund*, est responsable de la spécification des yeux chez les Bilatériens. Avant de commencer notre travail, une expression spécifique de *LsPax-6* au niveau des yeux en développement ou en cours de régénération était déjà connue chez *Lineus*. De plus, comme nous avions déjà décrit que l'inactivation de *LsPax-6* par ARN interférence (ARNi) chez l'adulte *L.sanguineus* conduisait à la "disparition" des yeux, nous savions que l'expression de *LsPax-6* est nécessaire au maintien des yeux. Afin de poursuivre notre étude, nous avons entrepris de tester l'implication éventuelle d'autres membres du "réseau génétique de détermination rétinienne" dans la spécification de l'oeil chez *Lineus*. Nous avons tout d'abord rapporté l'existence de trois gènes *Six* dans le génome de *L.sanguineus*: *LsSix1/2*, *LsSix3/6* et *LsSix4/5*. Les résultats de l'étude de leur expressions ont permis de laisser présager de leur implication générale dans le développement et la régénération de nombreux organes sensoriels, et notamment des yeux (spécialement dans le cas de *LsSix1/2*). Nous avons ensuite montré qu'un marquage par un anticorps anti-Dac permet d'envisager une possible expression d'un homologue de *Dachshund* au niveau des yeux de *Lineus* en développement. Ainsi, des facteurs clés, membres conservés du réseau de détermination rétinienne et iniateurs d'une cascade génétique conduisant à l'édification d'un oeil chez de nombreux Bilatériens, sont également impliqués dans la formation et le maintien des yeux chez *Lineus*. Cette unité moléculaire qui sous-tend, au sein des Bilatériens, la spécification de l'oeil, plaide en faveur de l'hypothèse d'une origine monophylétique de l'oeil.

Acknowledgments

I am deeply grateful to Prof. Walter J. Gehring for giving me the opportunity to perform my Ph.D. thesis in his laboratory. I would like to thank him for the interest he always had in my work and his constant enthusiasm. His support, advices and new ideas greatly helped me and encouraged me when experiments were not working as well as I would have liked. The stimulating atmosphere he created in his group by bringing together people from different background working on several different but complementary projects inspired me.

I am also deeply grateful to Prof. Jacques Bierne and Dr. Michel Tarpin for giving me the opportunity to perform my Ph.D. work on a subject that has been developed throughout many years in their laboratory in Reims, France. I would like to thank them for their support, advices and for the time spent to discuss my results on weekends, when I was coming back to Reims.

I would like to thank especially Dr. Marie Kmita and Prof. Heinrich Reichert who have accepted to review my thesis manuscript and be member of my thesis committee.

Furthermore, I would like to thank all the members of the group who contributed in a way or in another to my work with their help, experience and advices. Among them, I want to thank Paul Baumgartner who introduced me to molecular biology when I arrived in the laboratory to perform my diploma work. I also want to thank Lydia Michaut who helped me a lot with RNA extraction, mRNA quality checking and RT-PCR procedures. I would like to thank Urs Kloter and Nicole Grieder who always found some times to discuss my *in situ* protocol problem with me. I went through several months of nightmares with *in situ* troubleshooting, your advices really helped me to stay motivated when nothing at all was working! I also want to thank Hiroshi Suga for its help with phylogenetic analysis and scientific discussion, Makiko Seimiya for advices and scientific discussion. Both of them, in addition to Yoshi Adachi, know that I am in love with Japanese food, and especially with “yokan”: they often supplied me with “Japanese specialities” directly imported from Japan! Thanks a lot! I am definitively convinced I should go to visit Japan! I also want to thank Lukas Keller and Robin Vuilleumier for scientific discussions. It

was nice to start our Ph.D thesis at the same time! Especially to start it with working together for two months in the station biologique de Banyuls: working all July and August in the "small Banyuls' lab" of Walter was rather nice, don't you think?! And also Marina Beaufils, who shared the office with me: "Voilà!". Many thanks to Greta Backhaus for her secretarial assistance and for being so helpful. I also would like to thank Karin Mauro, Bernadette Bruno and Gina Evora, who support our scientific work by supplying buffers and by cleaning all the lab goods.

Many thanks to Nicole Grieder and Makiko Seimiya for comments and critical reading of this thesis manuscript. I know it was sometimes challenging to understand my "frenchenglish"!

During my time in Basel, I met several people who became close friends. I hope we'll stay in contact and will always enjoy a glass of wine …or champagne! together: Cristina, you know I'll always remember our "cDNA library" prep…actually our "separation column" is still on the wall of the lab! Daria, my "official fitness partner"! Nicole, who is as addicted as I am to coffee. Caroline, who shares my passion for dancing. And of course all the members of the Biozentrum 2nd floor "French connection"!

I enjoyed a lot my every-year "Roscoff retreat". Even though it was often painfull to collect *Lineus* in Roscoff during winter as it was amazingly windy and cold, eating "crêpes bretonnes" and oysters for three days helped me to recover well everytime!! I would like to thank Mme Sanséau who organized my stay at the "station biologique de Roscoff" every year. I also would like to thank the owner of "la crêperie les amours jaunes": I enjoyed a lot our discussion about books! I was waiting for our "rentrée littéraire" every year!

I also would like to thank my lovely "lulu", my parents and old-friends for all their love and constant support throughout all my life. Last, but not least, many thanks to Bertrand for his love, support and for being so patient during the writing time of this manuscript.

This thesis work was performed between two laboratories:

Laboratoire de Biologie Cellulaire et Moléculaire
Université de Reims Champagne-Ardennne
UFR Sciences Exactes et Naturelles
Moulin de la House BP1039
FR-51687 Reims Cedex 2/ France

Laboratory of Prof. Gehring
Focal area Growth & Development
Biozentrum, University of Basel
Klingelbergstrasse 50/70
CH-4056 Basel / Switzerland

Table of Contents

Short summary 3
Résumé court 4
Abstract 5
Résumé 7
Acknowledgments 9
Table of Contents 12
Abbreviations 17
List of figures 20

I. General Introduction 22

1. When evolutionary biology meets developmental biology 23
2. Choice of model organisms: the lophotrochozoan members as emerging systems 23
3. The homeobox-containing genes, a central focus for the body plan formation 26
3.1 Discovery of the homeobox 26
3.2 Structure of the homeodomain 27
3.3 The homeodomain, an efficient multifunction domain 29
3.4 The homeobox gene superfamily 30
3.5 The *Hox* genes 31
3.5.1 The *Hox* spatial and temporal colinearity rules 33
3.5.2 The Hox phenotypic suppression 34
3.5.3 The Hox co-factors 35
3.5.4 *Hox* genes and morphological innovations during evolution 36
3.5.5 Evolution of the *Hox* cluster 36
3.5.6 Evolutionary constraints on *Hox* genes clustering 38
3.5.7 Origin of the *Hox* cluster 38
3.5.8 The *Hox*, *ParaHox* and *NK* cluster: the "megacluster hypothesis" 39
4. Nemerteans 40
4.1 General characteristics of nemerteans 41
4.2 Regeneration and developmental plasticity in *Lineus* 43
4.3. *LsHox* and *LsPax-6* homologues 44
5. Why working with *Lineus* 46
6. References of the general introduction 48

II. Involvement of *Ls-Otx* and *Ls-Cdx*, two homeobox-containing genes, in the antero-posterior patterning of the body axis during both, development and regeneration of *L.sanguineus* 58

1. Introduction 59
1.1 The *Otx*/ *Otd* related genes 59
1.1.1 The *Drosophila Otd* gene 60

1.1.2 The mouse *Otx* genes 60
1.1.3 Otx-related gain-of-funtion 61
1.1.4 *Otx* gene cross-phylum rescue experiment 61
1.2 The *Cdx*/ *cad* related genes 63
1.2.1 The *Drosophila cad* gene 64
1.2.2 The vertebrate *Cdx* homologues 65
1.2.3 *Cdx*-related genes and the axial elongation process 66
1.3 *Otx*-related and *Cdx*-related genes, two ancestral key players of early A-P patterning in bilaterian animals? 67
1.3.1 Evidence from model organism studies 67
1.3.2 Evidence from *Nasonia* and *Tribolium* studies 69
1.4 *Otx*-related genes and the CNS development 72
1.4.1 Apparent dissimilarities in the CNS development of invertebrates and vertebrates 72
1.4.2 Homologous genes are involved in the CNS development of both invertebrates and vertebrates 73
1.4.3 The tripartite brain organization hypothesis 76

2. Aim of the first project of this PhD thesis 79

3. Results 81
3.1 Identification of *Otx* and *Cdx* homologues from *Lineus sanguineus* 81
3.1.1 Cloning of *Ls-Otx* 81
3.1.2 Cloning of *Ls-Cdx* 83
3.2 Expression patterns of *Lineus Otx* and *Cdx* homologues during development 84
3.2.1 The use of *L.viridis* to investigate *Lineus* development 84
3.2.2 *Otx* expression pattern during *Lineus* development 85
3.2.3 *Cdx* expression pattern during *Lineus* development 86
3.3 Expression patterns of *Ls-Otx* and *Ls-Cdx* in adults and during regeneration 87
3.3.1 *Ls-Otx* expression in adults *L.sanguineus* 87
3.3.2 *Ls-Cdx* expression in adults *L.sanguineus* 88
3.3.3. *Ls-Otx* expression during regeneration of *L.sanguineus* antecerebral end 89
3.3.4 *Ls-Cdx* expression during posterior regeneration of the *L.sanguineus* gut 90
3.3 Real time PCR approach 91
3.3.1 General considerations of the real time PCR procedure 91
3.3.2 Analysis *Ls-Otx* and *Ls-Cdx* expression levels by real-time PCR in adult *L.sanguineus* 92
3.4 Analysis of *Ls-Otx* and *Ls-Cdx* expression during both, anterior and posterior regeneration of an isolated fragment 8 94
3.4.1 ISH approach 95
3.4.2 Real-time PCR approach 99
3.5 Analysis of *Ls-Otx* and *Ls-Cdx* expression during anterior and posterior regeneration of worms missing body regions 1 to 5 and worms missing body regions 5 to 10, respectively 101
3.5.1 ISH approach 102

3.5.1.1 *Ls-Otx* and *Ls-Cdx* expression during anterior regeneration from worms missing body regions 1 to 5: "bottom parts" ……..… 103
3.5.1.2 *Ls-Otx* and *Ls-Cdx* expression during posterior regeneration from worms missing body regions 6 to 10 "upper parts" …… 108
3.5.2 Real-time PCR approach ……..… 111
3.6 Investigation of *Otx* expression in *L.lacteus*, a close relative of *L.sanguineus*, which exhibits different regeneration properties ……...….... 113

4. DISCUSSION …… 117
4.1 Isolation and characterization of *Ls-Otx* and *Ls-Cdx* ……. 117
4.1.1 The *Ls-Otx* gene…….…….. 117
4.1.2 The *Ls-Cdx* gene …….. 118
4.2 The most anterior and posterior *Lineus* body regions are specificied by *Otx* and *Cdx* homologues, respectively ……...…… 119
4.3 *Ls-Cdx* is specifically involved in the posterior regeneration …… 124
4.4 *Otx* displays dual roles in *L.sanguineus* regeneration …… 125
4.4.1 General involvement of *Ls-Otx* in the onset of regeneration .. 126
4.4.2 Specific involvement of *Ls-Otx* in anterior regeneration …… 127
4.5 Development and regeneration use similar molecular mechanisms but not always identical ones ……..…. 127
4.6 The different regeneration capacities of *L.sanguineus* and *L.lacteus* are reflected in differences in the expression pattern of *Otx* after amputation .. 128

5. References of chapter II ……..…….. 132

III. Involvement of some evolutionarily conserved key players of the "Retinal Determination Genetic Network" in the development, maintenance and regeneration of the eyes of *L.sanguineus* 143

1. Introduction …… 144
1.1 The perception of light ……...…… 144
1.2 Origin of the eyes ……... 146
1.3 Similarities and differences in various eye-types ……..…….. 148
1.3.1 Structure and developmental mode of the *Drosophila* compound eye *vs* the vertebrate eye ……..…..... 148
1.3.2 The ultrastructure of photoreceptor cells ……..……. 151
1.3.3 Photopigments ……..… 151
1.3.4 The extremely sensitive rhodopsin molecule …….……. 152
1.3.4.1 The phototransduction cascade …… 152
1.3.4.2 Rhodopsin regeneration after photobleaching …..... 154
1.3.5 Single origin of photoreceptor cells ……. 155
1.4 The retinal determination gene network: the RDGN ……..... 156
1.4.1 The Pax-6 transcription factor ……. 158
1.4.2 The Eya transcription factors ……..…… 160
1.4.3 The Six family members ……. 161
1.4.4 The Dach transcription factors …… 163

2. Aim of the second project of this PhD thesis ……...……... 165

3. Results …….. 167
3.1 *LsPax* genes ……..... 167
3.1.1 Characterization of the full-length *LsPax-6* and identification of a *LsPax-6* splice variant ……... 167
3.1.2. Identification of a *LsPax-2/5/8* gene ……..... 171
3.2. Isolation of a *Ls-opsin* gene and characterization of its expression pattern ……... 178
3.3 the *Ls-Six* genes …….. 180
3.3.1 Cloning and characterization of the *Ls-Six* genes ……...… 180
3.3.2 *Ls-Six1/2* expression patterns …….. 186
3.3.3 *Ls-Six3/6* expression patterns …….. 189
3.3.4 *Ls-Six4/5* expression patterns …….. 190
3.4 Cross-hybridization of an antibody anti-Dachshund with *Lineus* eyes .. 191

4. Discussion ……...……. 192
4.1 Conservation of a PST domain at the C-terminus of LsPax-6 ……... 192
4.2 A *LsPax-6* mRNA variant encodes a truncated LsPax-6 form …….. 193
4.3 Commonly accepted properties of the *Pax-2/5/8*-type HD are challenged by lophotrochozoan data ……...……... 197
4.4 Identification of an unusual *Ls-Opsin* specifically expressed during eye development ……. 199
4.5 First identification of homologues from the three *Six* gene subfamilies in a lophotrochozoan genome ……. 200
4.6 General involvement of the three *Ls-Six* genes in CNS and anterior sensory organ development …….. 201
4.7 Some members of the Retinal Determination Genetic Network are co-expressed in developing, regenerating and differentiated *Lineus* eyes……. 205

5. References of chapter III ……...…… 210

IV. Materials and Methods …… 224

1. mRNA extraction ……...…….. 224
2. Reverse transcription ……... 224
3. Degenerated PCR ……. 224
4. Cloning of PCR products ……. 225
5. DNA sequencing …….. 225
6. 3’ and 5’ RACE PCR ……... 225
7. Gene expression level analysis by real-time PCR …….. 226
8. Preparation of RNA probes …….. 226
9. *in situ* hybridization protocol ..…… 227
10. Antibody staining ……... 229
11. Experimental animals ……. 230

V. Appendix …… 231

1. Phylogenetic tree showing *Ls-Otx* and *Ls-Cdx* genes in relation to other homeobox-containing genes ……...…… 231
2. The *Ls-Pax-6* sequence, which was available before this PhD work, and the deduced amino acid sequence of its putative ORF…..…….. 232
3. Comparison of the *Ls-Pax6* splice variant sequence with the canonical one ….. 234
4. Sequence comparison of the protein encoded by the canonical *LsPax-6* mRNA splice variant with the protein encoded by the *LsPax-6* alternative splice ……..... 234
5. Fragment of the genomic sequence of *LsPax-6* ……... 235
6. Phylogenetic tree showing the *Ls-opsin* in relation to other invertebrate and vertebrate-type opsins ……..…… 236
7. Alignment of the *Lv-Six1/2* and *Lv-Six3/6* fragments obtained by degenerated PCR and the corresponding homeobox sequences of the *Ls-Six1/2* and *Ls-Six3/6* genes..237
8. The *Ls-Bmp2/4* gene and the deduced amino acid sequence of its putative ORF……238
9. The *Ls-Engrailed* gene and the deduced amino acid sequence of its putative ORF ……..……..239
10. The *Ls-Msx* gene and the deduced amino acid sequence of its putative ORF 240
11. Partial *Ls-Snail* gene and the deduced amino acid sequence of its putative ORF……. ……. 241
12. Partial *Ls-Twist* gene and the deduced amino acid sequence of its putative ORF……...……...…….……241

Abbreviations

aa: amino acid(s)

ANT-C: Antennapedia complex

Antp: Antennapedia

A-P: antero-posterior

Bcd: Bicoid

bHLH: basic helix-loop-helix

bp: base pair(s)

BrdU: bromodeoxyuridine

BX-C: Bithorax complex

Cad: Caudal

CNS: Central Nervous System

c-opsin: ciliary-type opsin

Dac: *Drosophila* Dachshund

Dach: vertebrate Dachshund

DCER: discoid ciliated epithelial region

DNA: desoxyribonucleic acid, cDNA: complementary DNA

dpa: days post amputation

Dpp: Decapentaplegic

dpt: days post transection

DTB: deutocerebral-tritocerebral boundary

D-V: dorso-ventral

ED: Eya domain

EF1α: Elongation factor 1 alpha

EGFR: epidermal growth factor receptor

En: Engrailed

Evo-devo: evolutionary developmental biology

Exd: Extradenticle

Ey: Eyeless

Eya: Eyes absent

Eyg: Eyegone

FGF: Fibroblast Growth Factor

GMP: guanosine monophosphate, cGMP: cyclic GMP

G-protein: GTP-binding protein

GPCR: G-protein coupled receptors

GRO: Groucho

GTP: guanosine triphosphate

Hb: hunchback

HD: homeodomain

Hh: Hedgehog

Hth: Homothorax

IP_3: inositol tri-phosphate

ISH: *in situ* hybridization

LNA: Locked Nucleic Acid

MHB: midbrain-hindbrain boundary

NMR: Nuclear Magnetic Resonance

ORF: Open Reading Frame

Otp: Orthopedia

Otd: Orthodenticle

PCR: Polymerase Chain Reaction

PD: Paired Domain

PDE: phosphodiesterase

PFA: paraformaldehyde

PIP_2: phosphatidyl inositol diphosphate

PLC: phospholipase C enzyme

PRD: paired-class

PST domain: Proline/ Serine/ Threonine-rich domain

RA: retinoic acid

RACE PCR: Rapid Amplification of cDNA ends PCR

RDGN: retinal determination genetic network

RK: Rhodopsin Kinase

RNA: ribonucleic acid, dsRNA: double-stranded RNA, mRNA: messenger RNA, rRNA: ribosomal RNA

RNAi: RNA-mediated gene interference

r-opsin: rhabdomeric-type opsin

SD: Six domain

So: Sine oculis

Sog: Short gastrulation

TALE: three amino acids loop extension

7TM domain: seven-transmembrane domain

Toy: Twin of eyeless

UTR: untranslated region

UV: Ultra Violet

Wg: Wingless

WSP motif: tryptophane/ serine/ proline-rich motif

Zen: Zerknüllt

List of figures

1.1 Comparison of homeodomain sequences encoded by the orthologous groups of *Hox* / *ParaHox* genes shared between nemertean (Lophotrochozoa), arthropods (Ecdysozoa) and chordates (Deuterostomia) **p.24**
1.2 The new molecule-based metazoan phylogeny **p.25**
1.3 Functional domains within the homeodomain required for homeoprotein intercellular transfer **p.30**
1.4 Classification of the homeobox genes based on phylogenetic analysis and chromosomal mapping for the ANTP class **p.31**
1.5 Chromosomal organization of a *Hox* cluster and schematic expression pattern in *Drosophila* and mouse **p.33**
1.6 Structure of the common ancestor of protostomian and deuterostomian Hox cluster, inferred from the structures of the insect and mammalian *Hox* clusters **p.37**
1.7 Scheme of *Lineus* body organization **p.42**
1.8 Paraffin section of a *Lineus ruber* eye **p.43**
1.9 Hypothetical evolution of the *L.sanguineus Hox* cluster **p.45**

2.1 *otd*/ *Otx* cross-phylum rescue experiments in *Drosophila* and the mouse **p.62**
2.2 Cdx members and the genetic network driving axial elongation, A-P patterning and mesoderm segmentation **p.66**
2.3 Morphogenesis of the nerve strand in prototypic Gastroneuralia and in prototypic Notoneuralia **p.73**
2.4 Representation of the molecular patterning of the DV axis in vertebrates and arthropods **p.75**
2.5 The tripartite organization of the *Drosophila* and mouse brain based on the expression patterns of orthologous genes **p.77**
2.6 The *Ls-Otx* ORF and its deduced aminoacid sequence **p.82**
2.7 Alignment of the *Lv-Otx* PCR fragment obtained by degenerated PCR and the *Ls-Otx* homeobox corresponding sequence **p.83**
2.8 The *Ls-Cdx* ORF and its deduced aminoacid sequence **p.84**
2.9 Expression analysis of *Ls-Otx* in developing *L.viridis* worms **p.86**
2.10 Expression analysis of *Ls-Cdx* in developing *L.viridis* worms **p.87**
2.11 *Ls-Otx* expression in an adult *L.sanguineus* head **p.88**
2.12 *Ls-Otx* expression during anterior regeneration of *L.sanguineus* brain **p.89**
2.13 *Ls-Cdx* expression during posterior regeneration of *L.sanguineus* gut **p.91**
2.14 *Ls-Otx* and *Ls-Cdx* expression levels measured by real-time PCR of different *L.sanguineus* body parts **p.94**
2.15 *Ls-Otx* and *Ls-Cdx* expression in isolated body fragment 8 undergoing both anterior and posterior regeneration **p.97**
2.16 Double ISH with *Ls-Otx* and *Ls-Cdx* anti-sense RNA probes on a regenerating isolated intestinal fragment, 8 days post-transection **p.98**
2.17 Expression levels of *Ls-Otx* and *Ls-Cdx* during both anterior and posterior regeneration of isolated intestinal fragment deduced by real-time PCR **p.99**
2.18 Adult *L.sanguineus* amputation scheme **p.102**
2.19 One-week-old *L.sanguineus* somatic embryo: details of the histological structure of the healing regions **p.104**
2.20 Differentiation of the DCER during somatic embryonic development **p.105**
2.21 *Ls-Otx* and *Ls-Cdx* expression during anterior regeneration from *L.sanguineus* lacking body regions 1 to 5 **p.107**
2.22 Posterior regeneration from adult *L.sanguineus* lacking body regions 6 to 10 is epimorphic **p.109**
2.23 *Ls-Otx* and *Ls-Cdx* expression during posterior regeneration from *L.sanguineus* lacking body regions 6 to 10 **p.110**
2.24 real-time PCR expression of *Ls-Otx* and *Ls-Cdx* during posterior regeneration of *L.sanguineus* amputated from body regions 6 to 10 **p.112**
2.25 The different regeneration properties of *L.sanguineus* and *L.lacteus* **p.114**
2.26 *Ll-Otx* detection in amputated *L.lacteus* **p.115**
2.27 *Ll-Otx* expression pattern at the plane of amputation **p.115**
2.28 *Otx* expression pattern in *L.sanguineus* during anterior regeneration **p.116**
2.29 Scheme of the grafting experiment that leads to regeneration of a complete worm form very few adult *L.sanguineus* cells originating from the antecerebral end **p.130**

3.1 Simplified illustrations of the building-plans of four types of eye **p.145**
3.2 The adult head of *Drosophila* develops from a pair of composite discs called the eye-antennal discs **p.148**
3.3 Development of the vertebrate eye **p.149**
3.4 The Mexican wave in eye development **p.150**
3.5 The two types of photoreceptors: the rhadomeric type and the ciliary type **p.153**
3.6 The different recovery paths of the rhabdomeric and ciliary visual pigments **p.154**
3.7 Eye specification genes in *Drosophila* **p.157**
3.8 Domain structures of the *Drosophila* RDGN key members **p.157**
3.9 Structural domains present in the different Pax proteins **p.159**
3.10 The full-length *LsPax-6* sequence and the deduced amino acid sequence of its putative ORF **p.168**
3.11 Scheme representing the *LsPax-6* alternative splicing process **p.169**
3.12 The deduced amino acid sequence of the putative ORF of the alternative splice **p.170**
3.13 The full-length *LsPax-2/5/8* sequence and the deduced amino acid sequence of its putative ORF **p.172**
3.14 Comparison of the LsPax-2/5/8 PD with the LsPax-6 PD **p.172**
3.15 Alignment of LsPax-2/5/8 protein with other Pax family members **p.177**
3.16 The full-length *Ls-opsin* sequence and the deduced amino acid sequence of its putative ORF **p.179**
3.17 Expression analysis of *Lv-opsin* in a developing *L.viridis* larva **p.180**
3.18 Sequence alignments of the SDs and HDs of mouse Six proteins **p.181**
3.19 Phylogenetic tree showing *Ls-six* genes in relation to other members of the *Six* gene family **p.182**
3.20 The full-length *Ls-Six1/2* sequence and the deduced amino acid sequence of its putative ORF **p.183**
3.21 The full-length *Ls-Six3/6* sequence and the deduced amino acid sequence of its putative ORF **p.184**
3.22 The full-length *Ls-Six4/5* sequence and the deduced amino acid sequence of its putative ORF **p.185**
3.23 Expression analysis of *Lv-Six1/2* in a developing *L.viridis* larva **p.186**
3.24 Expression analysis of *Ls-Six1/2* in an adult *L.sanguineus* worm **p.187**
3.25 Expression analysis of *Ls-Six1/2* during anterior head regeneration of *L.sanguineus* **p.188**
3.26 Expression analysis of *Lv-Six3/6* in developing *L.viridis* larvae **p.189**
3.27 Expression analysis of *Ls-Six3/6* in adult *L.sanguineus* worm **p.190**
3.28 Expression analysis of *Lv-Six4/5* in developing *L.viridis* larvae **p.190**
3.29 Expression analysis of *Ls-Six4/5* in adult *L.sanguineus* worm **p.191**
3.30 Cross-reaction of a monoclonal antibody developed against *Drosophila* Dac with *L.viridis* larvae **p.191**

Appendix
1. Phylogenetic tree showing *Ls-Otx* and *Ls-Cdx* genes in relation to other homeobox-containing genes **p.231**
2. The *Ls-Pax-6* sequence, which was available before this PhD work, and the deduced amino acid sequence of its putative ORF **p.232**
3. Comparison of the *Ls-Pax6* splice variant sequence with the canonical one **p.233**
4. Sequence comparison of the protein encoded by the canonical *LsPax-6* mRNA splice variant with the protein encoded by the *LsPax-6* alternative splice **p.234**
5. Fragment of the genomic sequence of *LsPax-6* **p.235**
6. Phylogenetic tree showing the *Ls-opsin* in relation to other invertebrate and vertebrate-type opsins **p.236**
7. Alignment of the *Lv-Six1/2* and *Lv-Six3/6* fragments obtained by degenerated PCR and the corresponding homeobox sequences of the *Ls-Six1/2* and *Ls-Six3/6* genes **p.237**
8. The *Ls-Bmp2/4* gene and the deduced amino acid sequence of its putative ORF **p.238**
9. The *Ls-Engrailed* gene and the deduced amino acid sequence of its putative ORF **p.239**
10. The *Ls-Msx* gene and the deduced amino acid sequence of its putative ORF **p.240**
11. Partial *Ls-Snail* gene and the deduced amino acid sequence of its putative ORF **p.241**
12. Partial *Ls-Twist* gene and the deduced amino acid sequence of its putative ORF **p.241**

CHAPTER I

General Introduction

1. When evolutionary biology meets developmental biology

For more than one century, zoologists have systematically classified the various phyla and reconstructed evolutionary trees based on anatomical data (Brusca and Brusca, 1990). While traditional zoology has highlighted the tremendous diversity of animal body plans in nature, molecular embryology has revealed the conspicuous unity underlying animal development and body patterning. One of the most striking findings in developmental biology over the past century represents the discovery of a set of highly conserved genes throughout evolution, the so-called "*Hox* genes" (see p.20). Unexpectedly, these genes exhibit an extreme high degree of conservation in their sequences and in their expression patterns throughout the bilaterian animals (Fig. 1.1). At the time of their discovery, they have been thought to represent a kind of "Rosetta stone" for understanding the common body plan of all living animals. This idea has galvanized the "evo-devo" community, which combines studies from the field of evolution with studies from the field of developmental biology in order to search for the "developmental synthesis of evolution" (Gilbert, 2003). The extreme *Hox* gene conservation in patterning the antero-posterior (A-P) body axis has served to define the concept of the metazoan "zootype" (Slack *et al.*, 1993). Furthermore, the investigation of *Hox* genes conservation has been considered as a paradigm for asking how changes in the embryonic gene expression program might give rise to morphological evolution. Indeed, variations in the number of *Hox* genes, in their sequences and in their spatial and temporal regulations have been proposed as a mechanism for body plan evolution and diversification (for review see: Gellon and McGinnis, 1998; Wagner *et al.*, 2003)

2. Choice of model organisms: the lophotrochozoan members as emerging systems

Recent molecular phylogenies, based on 18S ribosomal sequences (Halanych *et al.*, 1995; Aguinaldo *et al.*, 1997) and a number of other sequences, including the ones of the *Hox* genes, have radically transformed the classification of bilaterian animals.

Group 1 *Hox* genes identity (similarity)

```
PNTGRTNFTNKQLTELEKEFHFNKYLTRARRIEIAAALGLNETQVKIWFQNRRMKQKKRM LsHox1 (ribbonworm)
N-S-------------------R-----------NT-Q-------------------V DmLab (fruitfly)          88 (93)%
--N------T------------Y--------V------N-------------------E AmphiHox1 (lancelet)     90 (95)%
--AV-----T----------------------V----S-Q-------------------E Hoxa1/HOXA1             88 (95)%
                                                              (mouse/human)
```

Group 3 *Hox* genes

```
PKRSRTAYTSAQLVELEKEFHFNRYLCRRRRIEMAALLNLSERQIKIWFQNRRMKYKKDQ LsHox3 (ribbonworm)
A--A------------------------P------N----------------------E- CsHox3 (spider)        91 (95)%
G--A------------------------P--V----M---T-----------------E- AmphiHox3 (lancelet)  88 (97)%
S--A------------------------P--V---N------------------------ Hoxb3/HOXB3           91 (95)%
                                                              (mouse/human)
```

Group 4 *Hox* genes

```
SKRSRTAYTRHQILELEKEFHFNRYLTRRRRIEIAHALDLSERQIKIWFQNRRMKWKKEH LsHox4 (ribbonworm)
P--Q-----------------Y--------------T-V-------------------DN DmDfd (fruitfly)       88 (95)%
T---------Q-V------------------------S-G-T-----------------DN AmphiHox4 (lancelet)  86 (97)%
P---------Q-V------------------------T-C-------------------D- Hoxd4/HOXD4          90 (95)%
                                                              (mouse/human)
```

Group 6 *Hox* genes

```
QKRTRQTYTRYQTLELEKEFHFNKYLTRRRRIEIAHALGLTERQIKIWFQNRRMKWKKEN LsHox6 (ribbonworm)
P----------------------R------------S------------------A---- AlftzHox (mite)        93 (97)%
K--G-------------------------K-------L------------------------ AmphiHox6 (lancelet) 93 (95)%
GR-G-------------------R-----------N--C------------------------ Hoxa6/HOXA6         90 (95)%
                                                              (mouse/human)
```

Group 7 *Hox* genes

```
RKRGRQTYTRYQTLELEKEFHFNKYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKEN LsHox7 (ribbonworm)
----------------------R------------------------------------- DmAntp (fruitfly)       98 (100)%
------------------------------------------------------------- AmphiHox7 (lancelet)  100 (100)%
----------------------R-------------------------------------H Hoxa7/HOXA7          96 (100)%
                                                              (mouse/human)
```

Group 9 *Hox* genes

```
TRKKRKPYTRYQTMVLENEFLTNSYITRQKRWEISCKLHLTERQVKVWFQNRRMKRKKLN LsHox9 (ribbonworm)
V-------SKF--LE--K---F-A-VSK-----LARN-N-------I--------N--NS DmAbdB (fruitfly)      65 (86)%
S----C----F--LE--K---Y-M-L--ER-Y---QHVN-------I--------M--MS AmphiHox9 (lancelet)  67 (85)%
-----C---K---LE--K---F-M-L--DR-Y-VARV-N-------I--------M--M- Hoxc9/HOXC9          68 (85)%
                                                              (mouse/human)
```

Group Cdx *ParaHox* genes

```
KDKYRVVYSDRQRLELEKEFHYSRYITINRKAELAKSLDLTERQIKIWFQNRRAKERKIN LsCdx (ribbonworm)
--------T-F--------YCT------R--S---QT-S-S---V-----------TS- DmCad (fruitfly)       78 (88)%
----------H---------YSNK----K--VQ--NE-G-S---V----------Q--MA AmphiCdx (lancelet)  75 (83)%
--------T-H-----------------R--S---AN-G-----V-------------V- Cdx1/CDX1            86 (93)%
                                                              (mouse/human)
```

Fig. 1.1 Comparison of homeodomain sequences encoded by the orthologous groups of *Hox* / *ParaHox* genes shared between nemertean (Lophotrochozoa), arthropods (Ecdysozoa) and chordates (Deuterostomia) (From J. Bierne)
Nemertean: ribbonworm, ***L**ineus **s**anguineus* (*Ls*). Arthropoda: fruitfly, ***D**rosophila **m**elanogaster* (*Dm*); mite, ***A**rchegozetes **l**ongisetosus* (*Al*)*;* spider, ***C**upiennius **s**alei* (*Cs*). Cephalochordata: lancelet, *Branchiostoma floridae* (***Amphi**oxus*: *Amphi*). Chordata, Vertebrata: mouse, *Mus musculus/* human*, Homo sapiens*. Dashes indicate amino acid identity at similar position between nemertine and arthropod or chordate sequences. Similar amino acids are in light-faced letters, whereas divergent ones are in bold type.

They are now assumed to form a monophyletic group clearly separated from sponges, cnidaria and ctenophores (Fig. 1.2). DNA sequence analyses subdivide the bilaterian animals into three superclades: the Deuterostomia, the Ecdysozoa and the Lophotrochozoa (Adoutte *et al.*, 2000). This new phylogeny implies that, for example, the body axes patterning systems used by arthropods and vertebrates have been inherited from a common ancestor of all bilaterian animals. This ancestral animal that gave rise to the Lophotrochozoa, the Ecdysozoa and the Deuterostomia, have been named Urbilateria (De Robertis and Sasai, 1996).

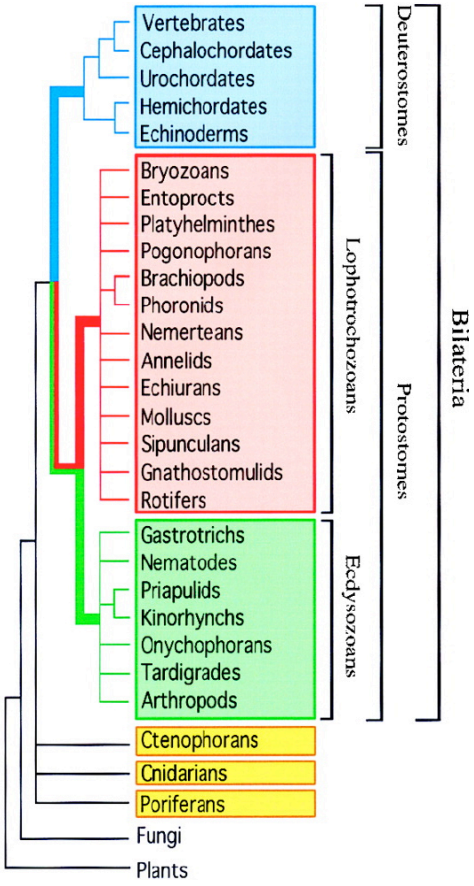


Fig. 1.2 The new molecule-based metazoan phylogeny (From Adoutte *et al.*, 2000)
A comparative approach was taken to achieve this metazoan phylogeny. Bilaterian animals undoubtedly form a monophyletic group, comprising the Deuterostomia, the Lophotrochozoa and the Ecdysozoa.

None of the major model organisms presently used belongs to the large branch of lophotrochozoan animals: *D.melanogaster* and *C.elegans* are ecdysozoan members, while *X.laevis, D.rerio* and *M.musculus* are deuterostomian animals. In addition, the ecdysozoan animals used as model organims are extremely derived, probably due to the ecological niches they occupy. Changes in the networks that control their

development must have occurred in order to restructure their body and to shorten their generation time. Indeed, many examples indicate that molecular networks of *Drosophila* and *Caenorhabditis* are characterized by sequence derivation and, importantly, loss of genes (Kortschak *et al.*, 2003; Tessmar-Raible and Arendt, 2003). As a consequence of the extensive gene loss that occured in the model invertebrate lineages, many genes only found in vertebrates have wrongly been considered as vertebrate-specific features. In fact, some of these genes can be traced back to pre-bilaterian times and thus must have been present in the common metazoan ancestor (Raible and Arendt, 2004; Kusserow *et al.*, 2005).

The lophotrochozoan members, constituting the third major branch of bilaterian animals, have received little attention, despite having the greatest diversity of the recognized body plans. Moreover, their developmental features, their organizations of the body plan and their gene structures are often considered as primitive, meaning, related to evolutionary ancestral conditions, or at least, less derived than the ones observed in the model organisms, currently used in the field of developmental biology. This is of particular interest for molecular comparisons among bilaterian members and therefore, for studying animal evolution. In addition, the interesting ability to regenerate and undergo asexual reproduction *via* fission, is a widespread lophotrochozoan characteristic, found in plathyhelmintes, annelids and nemerteans (reviewed in Sanchez-Alvarado, 2000). The investigation of regenerative processes, occurring in these relatively simple organisms, could bring light on various areas, such as tissue polarity, patterning and the control of size and proportions.

3. The homeobox-containing genes, a central focus for the body plan formation

3.1 Discovery of the homeobox

In order to explain the discontinuous jumps in body forms found during the evolution of species, naturalists have focused their attention on studying transformations of one body region into another one. In 1894, Bateson introduced the term “homeotic transformation”, in reference to the greek word homeosis, to describe phenotypic variation, in which “something is changed into the likeness of something

else". In 1915, Bridges and Morgan described the isolation of a fly with four wings, corresponding to the first homeotic transformation observed in *Drosophila* (Bridges and Morgan, 1923). This mutation, named *bithorax*, leads to the transformation of the third thoracic segment into a second thoracic segment. Consequently, the fly harbors an extra pair of wings (Bridges and Morgan, 1923). Lewis identified the gene responsible for the *bithorax* mutation. It is located on the right arm of the third chromosome, as part of a gene cluster (Lewis, 1978). That was the first discovery of a so-called homeotic gene, *HOM* or *Hox* gene. Using chromosome walking, the groups of Gehring and Scott headed for *Antp* and the *ANTC* (Garber et al., 1983; Scott et al., 1983). The cloning of the homeotic *Antennapedia* (*Antp*) gene led to the discovery of the homeobox, a 180 bp DNA segment characteristic for homeotic genes: indeed, when mapping the exons by hybridizing the cDNA clones to the chromosomal DNA isolated on the chromosome walk, cross-hybridization between *Antp* cDNA and a neighbouring gene was detected, which turned out to be *fushi tarazu* (*ftz*). This was the first sign of the homeobox (reviewed in Gehring, 1998). Cross-hybridizing sequences were subsequently found in the *Ultrabithorax* (*Ubx*) gene, and used to clone additional homeotic genes like *Deformed* (*Dfd*) and *abdominal A* (*abdA*) (McGinnis *et al.*, 1984a; Scott and Weiner, 1984). Fascinatingly, cross-hybridizations with an *Antp* homeobox fragment from *Drosophila* were also found with DNA from chicken, mouse, and human (McGinnis *et al.*, 1984b). Soon afterwards, the first vertebrate homeobox-containing gene was isolated by screening a *Xenopus* genomic library at low stringency with the homeobox fragment of the *Antp* gene from *Drosophila* (Carrasco *et al.*, 1984). Since then, over 1000 homeobox genes have been identified in many species (Abzhanov and Kaufman, 2000; Cook *et al.*, 2004; Levine *et al.*, 1984; Manuel *et al.*, 2006; Shepherd *et al.*, 1984; Sommer *et al.*, 1990; Seimiya *et al.*, 1994).

3.2 Structure of the homeodomain

Homeobox-containing genes have been found in the vast majority of today's existing phyla, from yeast, plants, sponges to humans (Levine *et al.*, 1984; Shepherd *et al.*, 1984; Sommer *et al.*, 1990; Seimiya *et al.*, 1994). The homeobox encodes the homeodomain (HD), generally composed of 60 amino acids. Its sequence is remarkably conserved throughout the metazoan kingdom: four hydrophobic core

amino acids, L16, F20, W48, F49 and three amino acids directly involved in the DNA binding R5, N51, R53 are identical in 95% of homeodomain sequences. Moreover, in more than 80% of the sequences, ten other amino acids are identical and twelve other positions of the HD vary only between two amino acids (Gehring *et al.*, 1994a). These conserved amino acids define the HD. They are crucial for the maintenance of the homeodomain 3D structure and for its DNA binding capacities.

The 3D homeodomain structure has been resolved by nuclear magnetic resonance (NMR) spectroscopy studies using the *Drosophila Antp* gene (Qian *et al.*, 1989). The HD folds into a compact globular structure, consisting of three α-helices. Helices I and II are arranged in an antiparallel manner relative to each other, while the third helix is arranged perpendicularly to the first two helices. Helix I is connected to helix II via a loop. Helix II, in association with helix III, forms a helix-turn-helix motif, a common motif present in many prokaryotic transcription factors. A flexible N-terminal arm, which is involved in DNA binding and in protein-protein interaction, precedes the first helix (Qian *et al.*, 1989). The HD binds with high affinity to its DNA-binding site (Affolter *et al.*, 1990). A common model for homeodomain-DNA complexes has been put forward based on NMR spectroscopy and crystallography studies (Kissinger *et al.*, 1990; Otting *et al.*, 1990; Billeter *et al.*, 1993). The helix III, the so-called "recognition helix", makes contact with specific bases in the major groove of the DNA. Functional analyses of several homeodomains have revealed the crucial importance of the amino acid at position 50 in the recognition helix (Treisman *et al.*, 1989; Schier and Gehring, 1992). Helix III establishes specific DNA contacts with the core motif TAAT, found in Hox target binding sites. In contrast, the flexible N-terminal part of the HD interacts with the minor groove. The loop between helix I and helix II comes to lie along the DNA backbone on the other side of the major groove (reviewed in Gehring *et al.*, 1994b). HDs can bind to DNA as monomers or as homo- or hetero-dimers in a sequence-specific manner. Two types of interactions between DNA and the HD have been described: the HD establishes specific contacts with DNA by hydrogen bonds. The most evolutionary conserved specific contact is the one between N51 of the HD and the second A of the target sequence "TAAT" (Wolberger *et al.*, 1991). In addition, the HD establishes unspecific contacts with the DNA backbone by salt bridges, hydrogen bonds, hydrophobic interactions and van der Waals interactions in order to stabilize the HD-DNA complex.

3.3 The homeodomain, an efficient multifunction domain

The HD possesses a well-described activity as a transcription factor. It binds to DNA and acts as a transcriptional activator or repressor (Schier and Gehring, 1992; Zhang *et al.,* 1996; Li *et al.,* 1999). In addition, it also regulates translation. For instance, the Bicoid-HD can repress translation of *caudal* by binding to the 3' UTR of its mRNA and by interacting with a key regulator of translation: eIF4E, a factor involved in the initiation of translation (Rivera-Pomar *et al.*, 1996; Niessing *et al.*, 2000; Nedelec *et al.*, 2004). Importantly, the domain of interaction with eIF4E has been found in 200 other homeoproteins (Topisirovic *et al.*, 2003). Furthermore, the HD is also implicated in protein-protein interactions: for example, helices I and II of the Antp-HD are involved in the mutual inhibition of Eyeless (Ey) and Antp at the protein level (Plaza *et al.*, 2001). In addition, the HD is also involved in secretion and internalization of homeoproteins, allowing them to travel between cells (Fig. 1.3). There is some evidence for intercellular trafficking of several homeodomain-containing transcription factors such as Engrailed (En), some Hox, Emx1 and Emx2, Otx2 and Pax-6 (Joliot *et al.*, 1998; Prochiantz and Joliot, 2003). The recently proposed capacity for homeoproteins to travel between cells raises questions about its developmental implications. Some data support the idea that these homeoproteins could have non-autonomous cell activities, therefore paracrine functions. Hence, the possibility that transcription factors are at the basis of a new mode of signal transduction is now debated: since many homeoproteins contribute to positional information, their transfer from one cell to another would be an efficient and economic way to combine signal transduction and positional information (Prochiantz and Joliot, 2003). Last but not least, an external gradient of En-2, showing a morphogen-like activity, plays a pivotal role in axon guidance in the developing midbrain of *Xenopus* (Brunet *et al.*, 2005). The function of the two homeoproteins Otx2 and Gbx2 in the formation of the vertebrate brain compartments also argues for a possible morphogen-like activity of homeoproteins (Prochiantz and Joliot, 2003).

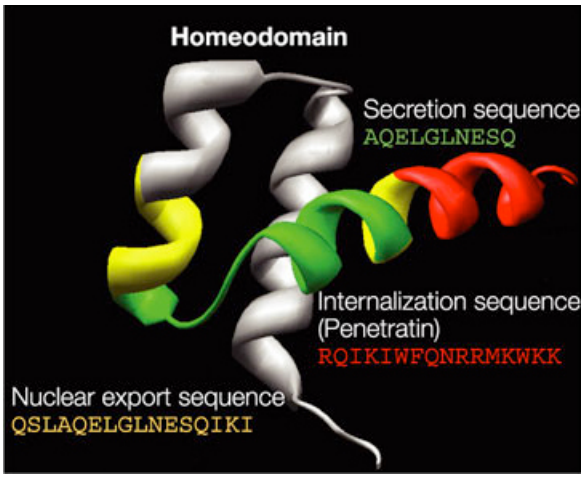


Fig. 1.3 Functional domains within the homeodomain required for homeoprotein intercellular transfer (From Prochiantz and Joliot, 2003)
Three domains that are required for secretion, internalization and nuclear export have been characterized by loss-of-function (deletion) or gain-of-function (synthetic peptides) studies. The secretion sequence (in green) is part of the nuclear export sequence (in yellow) and its deletion blocks nuclear export. The internalization sequence (in red) has been used as a vector to introduce cargo into living cells and is therefore also known as Penetratin. It is important to note that the secretion and internalization sequences are distinct.

3.4 The homeobox gene superfamily

All members of the homeobox gene superfamily are defined by the presence of a homeobox in their sequences. Many of them encode crucial transcription factors, which regulate the activity of other factors, such as signaling molecules. Homeoproteins play important roles at almost all levels of development. Their loss is typically associated with dramatic changes in the developmental program of many organisms, including humans (Hombria and Lovegrove, 2003). They are involved in numerous developmental events in all multicellular organisms, including the establishment of the A-P axis in the early embryo, eye development, heart development, striated muscle development, and formation of the head and the brain. Every homeobox sequence possesses some sequence specificities, also called "signatures", which have been used to classify homeobox-containing genes. They have been grouped into a number of distinct subfamilies based, not only on sequence similarities, but also on structural characteristics and phylogenetic analysis (Fig. 1.4).

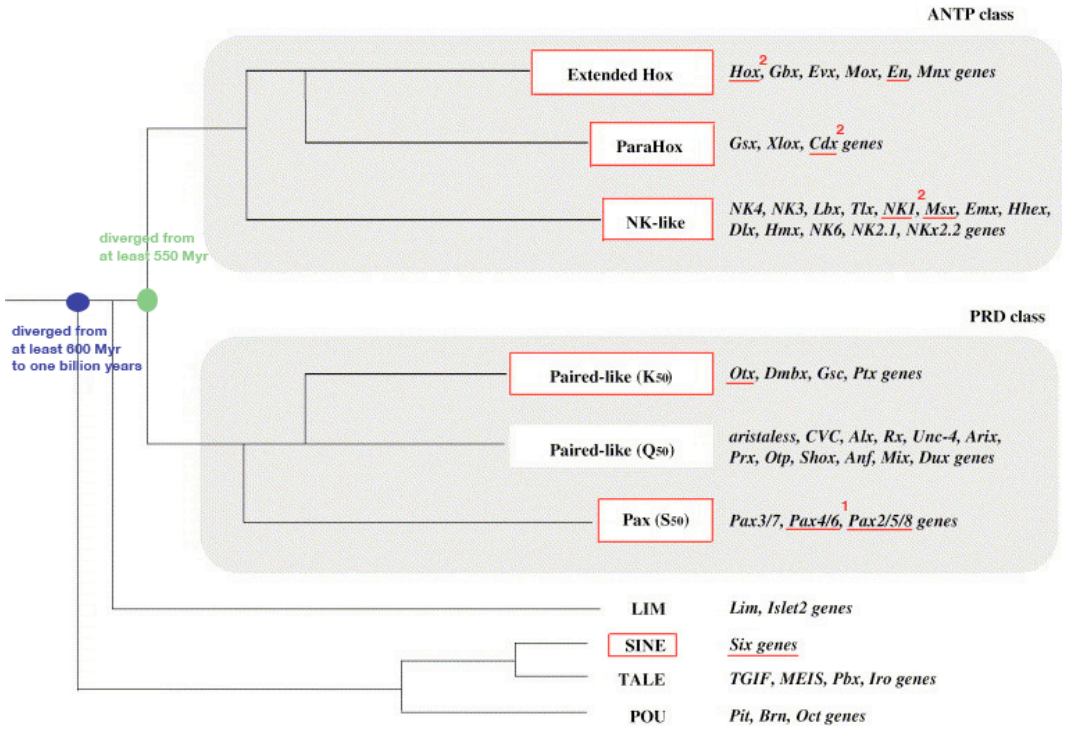


Fig. 1.4 Classification of the homeobox genes based on phylogenetic analysis and chromosomal mapping for the ANTP class (From Holland and Takahashi, 2005)
Homeobox genes are subdivided into two majors classes: ANTP and PRD plus several more divergent: LIM, SINE, TALE (they exhibit an unusual 63 aa long homeodomain) and POU. Representatives of the homologues found in the *Lineus sanguineus* genome are underlined in red. 1: identified by F. Loosli (Loosli *et al.,* 1996), 2: identified by M. Kmita (Kmita-Cunisse *et al.,* 1998), the unnumbered ones have been identified during this thesis work.

The major classes of homeobox-containing genes have diverged very early in eukaryotic evolution (Fig. 1.4): for example, members of the POU family have evolved independently from the ones of ANTP for at least 600 to 2000 M years. Members of ANTP and PRD classes have evolved later, since they are only found in metazoans, suggesting that they probably diverged around 550 M years ago.

3.5 The *Hox* genes

The *Hox* genes are the most famous members of the homeobox gene superfamily. They were the first gene family shown to act in similar and probably homologous ways in both insect and vertebrate development (Gehring, 1993). They are known to specify the body plan of multicellular organisms, by assigning different identities to cells along the A-P axis of the developing organisms (McGinnis and Krumlauf, 1992).

They have been found in most animal phyla, showing an extraordinary conservation of their homeodomain sequences and their expression patterns along the body axis as well as their chromosomal organization as clusters (Fig. 1.5). The *Drosophila melanogaster Hox* cluster is split into two: the Antennapedia complex (ANT-C) and the Bithorax complex (BX-C). ANT-C contains five *Hox* genes: *labial*, *Proboscipedia*, *Deformed*, *Sex combs reduced* and *Antennapedia* (Kaufman *et al.*, 1990), while BX-C contains three *Hox* genes: *Ultrabithorax*, *abdominal*-A and *Abdominal*-B (Sanchez-Herrero *et al.*, 1985). Two duplication events, which took place early in vertebrate evolution, have generated the four *Hox* clusters, seen in mammals and birds (Holland *et al.*, 1994). During evolution, some gene loss has occurred in each vertebrate cluster: therefore not every type of vertebrate *Hox* gene is represented in each of the four clusters. Each vertebrate cluster consists of 13 paralogue groups with nine to eleven members. These have been assigned on the basis of sequence similarity and relative position within the cluster (Duboule, 1992).

When misexpressed, the homeotic genes are able to induce transformations of the body plan. One of their key features is that loss- and gain-of-function mutations of these genes lead to opposite homeotic transformations. For example, in the originally called *Nasobemia* mutation, the fly antennae are transformed into a pair of middle legs (Gehring, 1966). Actually, this mutation corresponds to a dominant gain-of-function of *Antp*, which leads to transformations of the first thoracic segment (T1) and some head segments into the second thoracic segment (T2). In contrast, recessive loss-of-function of *Antp* transforms T2 into T1. As T2 is lacking in loss-of-function mutants and as an additional T2 is generated in gain-of-function mutants, this suggests that *Antp* specifies T2. Furthermore, there is an evident correlation between a shift in the anterior boundaries of *Hox* expression and changes in the identity of axial structures. This has led to the suggestion that the combinatorial distribution of Hox products in a specific region serves as an axial code, the Hox code (Kessel and Gruss, 1991; Kessel, 1992).

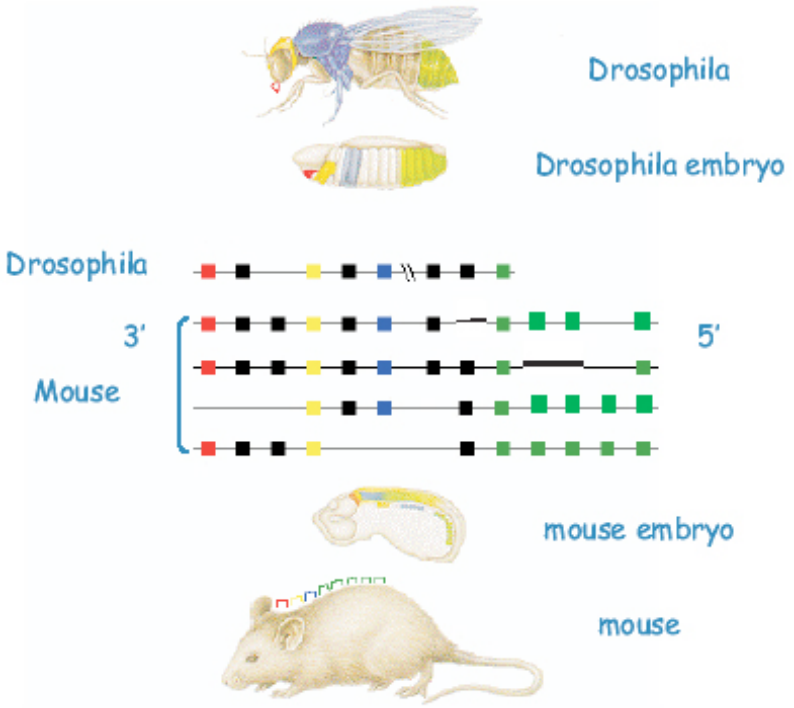


Fig. 1.5 Chromosomal organization of a *Hox* cluster and schematic expression pattern in *Drosophila* and mouse (From Stern *et al.*, 2006)
The fly has only one *Hox* cluster, while the mouse has four. *Hox* clusters are spatially and temporally collinear: genes located towards the 3' end of the cluster are expressed more rostrally and earlier than those closer to the 5' end.

3.5.1 The *Hox* spatial and temporal colinearity rules

As homeodomain-containing transcription factors, the *Hox* genes specify diverse body regions by the regulation of a unique set of downstream target genes, which triggers alternative developmental pathways (Akam, 1998). The *Hox* genes, which are expressed in specific combinations and specific concentrations in discrete domains of the A-P body axis, give unique positional information to the cells. Lewis has postulated that their physical arrangement along the chromosome control their physical order of expression along the A-P axis of the developing embryo: this correspondence has been called "the spatial colinearity rule" (Lewis, 1978). Firstly described in the fly, this correspondence between body axis specification and genomic organization is evolutionarily conserved in the homeotic clusters of most animals (McGinnis and Krumlauf, 1992). Furthermore, in mammals and short-germ band insects, unlike *Drosophila*, an additional phenomenon is observed: the *Hox* genes localized at the 3' end of the cluster are expressed first, whereas more 5' *Hox* genes are expressed later and sequentially. This is known as "the temporal colinearity

rule" (Kmita and Duboule, 2003). In mammals, the temporal colinearity is correlated with a modulation of the chromatin conformation of the *Hox* clusters (Chambeyron and Bickmore, 2004). Spatial and temporal colinearity can be either mechanistically linked, as the *Hox* axial limit of expression can be determined by the time of gene initiation (Duboule, 1994), or mechanistically independent, as different regulatory regions driving temporal or spatial colinearity have been found in the mouse genome (Tarchini and Duboule, 2006). Analysis of the molecular mechanisms, which underly both the spatial and temporal colinearity, suggests that various species-dependent processes exist to achieve the proper spatio-temporal expression pattern of *Hox* genes. Thus, it seems that strategies are not so important as long as the correct Hox protein distribution is accomplished (Kmita and Duboule, 2003).

3.5.2 The Hox phenotypic suppression

There is a functional hierarchy among homeotic gene members: in *Drosophila*, one homeotic protein can usually impose its function over more anterior co-expressed homeotic genes, through a suppressive mechanism called "phenotypic suppression" (Gonzalez-Reyes *et al.*, 1990). Thus, the general tendency is that one homeotic protein inactivates others that act anterior to its own domain, and is inactivated by others that act posterior to its domain. However, the homeotic genes do not always act in respect to a strict hierarchy, they also act in a combinatorial way (Duboule and Morata, 1994). A similar mechanism, named "posterior prevalence", is observed for the vertebrate *Hox* genes (Duboule, 1991). The suppressive mechanism of functional hierarchy found among the *Drosophila* Hox proteins and the vertebrates HOX ones does not seem to occur at the transcriptional or translational level (Schock *et al.*, 2000; Williams *et al.*, 2006). Some data rather suggest that this suppressive mechanism is based on protein-protein interactions in vertebrates (Williams *et al.*, 2006) and in the fly, as well (Y. Adachi, personal communication). Functional repressions of homeodomain-containing transcription factors by direct protein-protein interactions have already been described: for example, it has been demonstrated that the HD of Antp can interact with the paired domain (PD) and/ or the HD of Ey (Plaza *et al.*, 2001). Recent findings suggest that protein-protein interactions could be the molecular basis of the suppressive mechanism observed among both, the *Drosophila* homeotic genes and the vertebrates *Hox* genes (Williams *et al.*, 2006).

These protein-protein interactions can be mediated by direct HD interactions. However, it is not known whether this mechanism can be attributed solely to the amino acid sequence of the HD. Other non-homeodomain regions of both, *Drosophila* and vertebrate Hox proteins, contain co-factor interaction motifs and activation or repression domains (Yaron *et al.*, 2001; Merabet *et al.*, 2003; Williams *et al.*, 2005). This offers numerous possibilities of protein-protein interactions between Hox proteins, not only mediated by the HD. Indeed, it has been shown that the HD of vertebrate HOX group 13 is neither necessary nor sufficient for the occurrence of posterior prevalence in the mouse limb (Williams *et al.*, 2006).

3.5.3 The Hox co-factors

The *Hox* genes specify body region identity by activating and/ or repressing a unique set of target genes. However, the fact that most Hox proteins bind to similar and relatively simple DNA sequences *in vitro* and *in vivo* (Biggin and McGinnis, 1997), has raised the question of the achievement of their downstream target specificity and their ultimate initiation of distinct developmental pathways. One likely possibility is that other factors modulate Hox protein binding specificity, but only few Hox cofactors have been identified, so far (Mann and Morata, 2000). It is known that Hox proteins bind to DNA in association with TALE (three amino acid loop extension) homeodomain-containing proteins as co-factors: Extradenticle (Exd in Drosophila, Pbx in vertebrates) and Homothorax (Hth in Drosophila, Meis in vertebrates) (Moens and Selleri, 2006). This association with TALE homeodomain-containing co-factors improves the Hox-DNA binding selectivity. These TALE co-factors are required for a wide range of Hox functions, whereas a recently discovered novel class of Hox co-factors, assigned as “contextual Hox co-factors”, are required for only a subset of Hox protein activities (Merabet *et al.*, 2005). They recruit specific transcriptional co-activators or co-repressors. By being spatially restricted and selectively recruited by target-specific cis-regulatory sequences, the contextual Hox co-factors are thought to provide the cell-type and target-type specificity of Hox protein function (Merabet *et al.*, 2005).

3.5.4 *Hox* genes and morphological innovations during evolution

The *Hox* genes are able to induce body plan transformations when misexpressed. This capability has led to the assumption that duplication events and misexpression of *Hox* genes during evolution have been necessary for generating the observed morphological diversity found in metazoans. It is tempting to speculate that the recruitment of regulatory genes, or pre-existing regulatory networks has been a common mechanism for the generation of novel body structures during evolution (Lowe and Wray, 1997; Keys *et al.*, 1999). Changes in *Hox* gene expression have been correlated with key events in the evolution of the ecdysozoan body plan (Averof and Patel, 1997) and of the deuterostomian one (Burke *et al.*, 1995). Also in Lophotrochozoa, the bilaterian clade harboring the greatest body plan diversity, some data point at a correlation between specific *Hox* expression and morphological innovations (Lee *et al.*, 2003).

3.5.5 Evolution of the *Hox* cluster

Hox clusters are found in almost all bilaterian animals. The persistence of a cluster organization of *Hox* genes throughout bilaterian genomes indicates that this is likely due to evolutionary constraints (Hurst *et al.*, 2004). Evolutionary analyses suggest that the common ancestor of bilaterian animals probably possessed a single *Hox* cluster (De Rosa *et al.*, 1999). Comparisons of bilaterian *Hox* cluster have suggested that the urbilaterian *Hox* cluster could have been composed of seven *Hox* genes (Fig. 1.6): two genes from the anterior group, one group 3 gene, three genes from the central group and a single gene from the posterior group (Garcia-Fernandez, 2005a). Ecdysozoan and lophotrochozoan members possess only one *Hox* cluster in their genomes. In *Drosophila*, the complex is split into two. In addition, some non-*Hox* genes sit within the clusters (Sanchez-Herrero *et al.*, 1985; Kaufman *et al.*, 1990). The *Hox* clusters of nematodes are the results of extensive gene loss, associated with rapid sequence evolution (Aboobaker and Blaxter, 2003). *C.elegans* shows the most derived state of a nematode *Hox* cluster. It contains a greatly reduced numbers of *Hox* genes: only six *Hox* genes are present in its genome (Bürglin and Ruvkun, 1993).

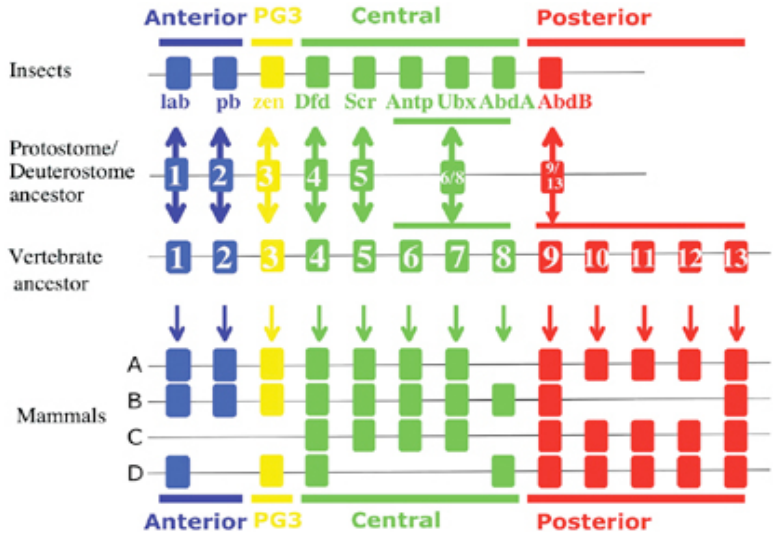


Fig. 1.6 Structure of the common ancestor of protostomian and deuterostomian Hox cluster, inferred from the structures of the insect and mammalian *Hox* clusters (From Garcia-Fernandez, 2005a)
Hox genes are grouped in anterior, group 3, central and posterior classes based on sequence similarities. Numbers and arrows indicate orthology relationships.

In contrast to protostomian genomes, the deuterostomian genomes contain different numbers of *Hox* clusters. Amphioxus has a single, intact *Hox* cluster of 14 genes (Garcia-Fernandez and Holland, 1994). This is the only deuterostomian *Hox* cluster observed so far, which is not only intact but also does not exhibit gene loss and has retained the *Hox* genes in their ancestral order. Within the lineage leading to vertebrates, gene duplications led to an expansion of the cluster and then, the cluster itself underwent duplications. Four copies of the cluster are found in human and mice genomes. In the case of the teleost fish, the *Hox* cluster has even been multiplied from seven to eight clusters (Crow *et al.*, 2006). In the non-vertebrate deuterostomian phylum urochordata, the central *Hox* genes have been lost and the ancestral *Hox* cluster has been fragmented: the nine *Hox* genes of *Ciona* present a disintegrated genomic organization. This correlates with the fact that there is only a remnant of spatial colinearity of the *Ciona Hox* gene expression (Ikuta *et al.*, 2004). The nine *Hox* genes from *Oikopleura*, another tunicate, are even more dispersed in its genome than what is observed in *Ciona*. Although *Oikopleura* has the most extreme case of *Hox* cluster disintegration, described so far, its *Hox* genes are expressed in staggered domains along the A-P axis, indicating that the spatial colinearity rule is retained (Seo *et al.*, 2004).

3.5.6 Evolutionary constraints on *Hox* genes clustering

Colinearity seems to be a widespread feature of *Hox* gene expression, and is often assumed to be a major reason for the evolutionary conserved cluster organization of the *Hox* genes. The several various derived lineages, which have a broken *Hox* cluster, such as the drosophilids, the nematodes and the urochordates, exhibit a rapid mode of embryogenesis. During the rapid development of such organisms, there is no opportunity for the *Hox* temporal colinearity to take place. Interestingly, the *Hox* spatial colinearity rule is retained (Monteiro and Ferrier, 2006). These observations suggest that the mechanisms generating spatial colinearity do not necessarily require *Hox* gene clustering in order to take place. They also suggest that the mechanisms generating temporal colinearity of the *Hox* gene expression may be the major constraining forces on *Hox* cluster organization (Patel, 2004). Indeed it has been shown that the clustered organization of *Hox* genes is required for the establishment of a tight temporal control of their expression, whereas this organization is dispensable for the correct spatial expression of the *Hox* genes in many other cases (Kmita and Duboule, 2003).

3.5.7 Origin of the *Hox* cluster

As the *Hox* genes display important roles during development of bilaterian animals, the origin of the *Hox* cluster has attracted considerable attention, but still remains unclear. The cnidarian phylum, which comprises sea anemones, corals, hydras and jellyfishes, is a sister group of the Bilateria. This non-bilaterian group represents a key transition in the evolution of animal complexity (Philippe *et al.*, 2005). Therefore, they can be critical for the investigation of the early history of homeobox-containing genes. *Hox*-like genes have been identified in many cnidarians (Gauchat *et al.*, 2000), but their status is often ambiguous. The consensus view is that a *Hox* cluster was already present in the ancestral cnidarian genome (Ferrier and Holland, 2001). During the development of the anthozoa sea anemone *Nematostella*, which is considered to represent the basal group within the cnidaria (Darling *et al.*, 2005), five *Hox* genes are expressed in staggered domains along the primary body axis (Finnerty *et al.*, 2004). It has been suggested that *Hox* genes are involved in patterning the primary body axis of cnidarian larvae (Finnerty *et al.*, 2004). Data from

another cnidaria, the hydrozoan jellyfish *Podocoryne*, are consistent with this view. However, the axial expression boundaries of *Podocoryne* homologues and *Nematostella* homologues at the larval stage are not conserved (Yanze *et al.*, 2001). Furthermore, linkages between distinct homeobox genes have been found in the *Nematostella* genome (Chourrout *et al.*, 2006). However, the existence within the cnidaria of a canonical *Hox* system, defined as a set of closely linked and interacting *Hox*-related genes responsible for patterning the A-P body axis, is a highly debated and controversial topic. Indeed, in disagreement with Chourrout and many others authors, Kamm *et al.* have argued that the cnidarian *Hox*-like genes do not conform to the *Hox* paradigm and that a true *Hox* system is absent in cnidaria, based on the analyses of *Hox* sequence relationships, gene organization and expression data (Kamm *et al.*, 2006). They have proposed that cnidaria split from the lineage leading to the Bilateria after the emergence of the ancestor of the *Antp* subclasses, but before the establishment of a canonical *Hox* system (Kamm *et al.*, 2006).

3.5.8 The *Hox*, *ParaHox* and *NK* clusters: the "megacluster hypothesis"

Two close relatives of the *Hox* cluster have been found in the recent years: the *ParaHox* cluster and the *NK* cluster. The *ParaHox* cluster is a paralogue of the *Hox* cluster: both *Hox* and *ParaHox* clusters arose by duplication of an ancestral *ProtoHox* cluster early in metazoan evolution (Brooke *et al.*, 1998). The *ProtoHox* cluster itself probably originated through *cis*-duplication of a founder *ProtoHox* gene (Garcia-Fernandez, 2005a). The *ParaHox* genes: *Caudal* (*Cad/ Cdx*), *genomic screened homeobox* (*Gsh*) and *Xenopus laevis homeobox 8* (*Xlox*) have been found in clusters only in chordates so far, even though the individual genes are found scattered in other deuterostomian and protostomian genomes (Ferrier and Minguillon, 2003). The second homeobox gene family, more distantly related, is the *NK-like* homeobox gene family. *NK-like* genes are found in clusters in some lineages (Luke *et al.*, 2003). The *Drosophila* genome contains a *NK* cluster of six genes, also called *93D/ E* cluster (Jagla *et al.*, 2001). The *NK* cluster has retained a compact organization in *Drosophila*, whereas it has been broken into three pieces in the chordate lineage. It has been proposed that a *NK* cluster of seven genes: *Msx-NK4-NK3-labdybird related homeobox (Lbx)- T-cell leukaemia homeobox (Tlx)- NK1-NK5*, was already present in the urbilaterian genome (Garcia-Fernandez, 2005b). Garcia-Fernandez has proposed a

model for the genesis and evolution of the *ANTP*-class homeobox genes: early in metazoan evolution, a *ProtoANTP* founder gene has generated two genes, a *ProtoHox*-like gene and a *ProtoNK* gene, by *cis*-duplication. Both, the *ProtoHox*-like and the *ProtoNK* genes, have been amplified by cis-duplication events. This has led to the generation of *Hox*, *ParaHox* and *NK* clusters, consisting of a "megacluster". Subsequent eventual chromosomal breakages could have split the megacluster and give rise to unlinked clusters (Garcia-Fernandez, 2005b).

4. Nemerteans

The phylum of nemerteans, also called ribbonworms, comprises more than 900 species of bilaterally symmetrical worms, which exhibit a flat, unsegmented body. Nemerteans have traditionally been considered as acoelomate animals (Gibson, 1972), based on some morphological similarities to the platyhelminthes, also known as planarians. However, structural analyses, supported by 18S rRNA sequence comparison, suggest that they are coelomate animals (Turbeville, 1991; Turbeville *et al.*, 1992). Furthermore, nemerteans have a body cavity with an eversible proboscis (rhynchocoel) and blood vessels that have been interpreted as coelomic cavities. In addition, they have a complete gut, which exhibits a separate ventral mouth communicating with the anterior foregut and a dorsal anal pore situated near the caudal end. In contrast, the platyhelminthes have only one opening in their digestive tract and no protosegmented structures, unlike the repeated nemertean gonads, intestinal coeca and cutaneous rings. Nemerteans occupy a basal position in the evolution of metazoans; they are clearly distinct from planaria (Carranza *et al.*, 1997), and their body plan may be close to the ancestral condition found in bilaterian animals. Nemerteans are classified as lophotrochozoan animals, among the Protostomia. Protostomia are defined as animals in which the blastopore gives rise to the mouth, whereas Deuterostomia are animals in which the mouth is formed secondarily, by a perforation of the ectoderm, opposite to the original blastopore; the anus arises at or close to the site of the original blastopore (De Robertis, 1997). Actually, this embryological distinction between Protostomia and Deuterostomia is not as evident as it seems, especially in the case of nemerteans: the position of the blastopore and the definitive mouth is intermediate between Proto- and Deuterostomia in different nemertean species (Nusbaum and Oxner, 1913; Iwata, 1985).

4.1 General characteristics of nemerteans

Nemerteans are mostly marine animals living in littoral or coastal regions. But some freshwater and terrestrial species have also been described. Most nemerteans are carnivorous: they show active predatory habits and feed on small animals, such as nematodes, crustaceans, annelids, mollusks and fish. Nemerteans move by using their epidermal cilia and their powerfull body wall musculature; some of them can actively swim. Their body size ranges from few millimeters to several meters in length, but the vast majority of them are less than twenty centimeters long. Nemerteans use their proboscis to capture prey, for self-defense and sometimes for locomotion. Nemerteans from the class Anopla, such as *Lineus*, have an unarmed proboscis, whereas the members of the Enopla class have an armed one, which carries a stylet at its anterior end. Nemerteans have a closed circulatory system, lined with an endothelium but do not have a heart. Most of them also possess a nephridial excretory system. Nemerteans have a lateral nervous system, rather than a ventral one as in most invertebrates or a dorsal one as in chordates: it consists of a pair of cerebral ganglia and two lateral nerve cords. The sense organs of nemerteans include the eyes, the cephalic grooves, the paired cerebral organs, some sensory epithelial cells and a frontal organ composed of ciliated cells located at the anterior end (Fig. 1.7).

Lineus, a marine nemertean from the class Anopla, has a distinct prototypical body organization along its A-P body axis: ten non-overlapping body regions can be distinguished from the rostral end to the caudal one. These body regions are defined based on specific morphological features (Fig. 1.7). For example, region 1 is characterized by the presence of eyes.

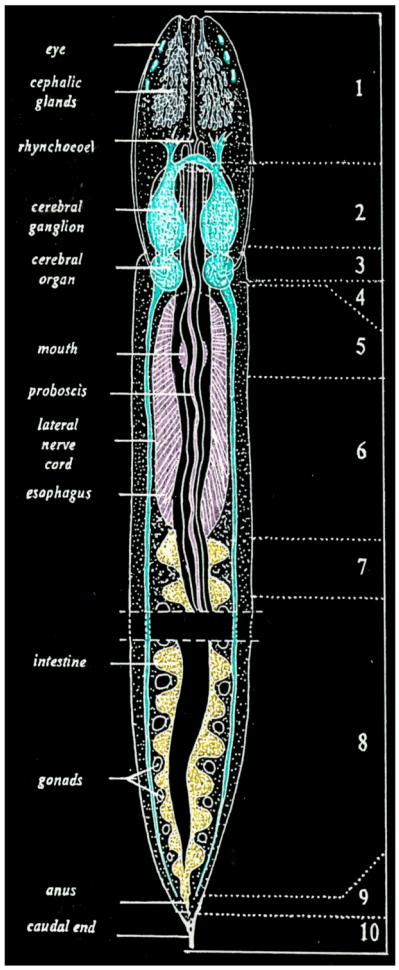


Fig. 1.7 Scheme of *Lineus* body organization
The different body regions are numbered starting from the anterior and are characterized by the following structures: the region "1": cephalic glands, rhyncodeum and eyes; region "2": cerebral ganglia; region "3": sensory cerebral organs; region "4": postcerebral and preesophageal connective tissue; region "5": ventral mouth; region "6": posterior esophagus and nephridia; region "7": anterior intestine (gonads are absent); region "8": middle and posterior intestines with serially repeated gonads; region "9": dorsal anal pore; region "10": caudal end.

The eyes of nemerteans are of the primitive pigment cup type and do not have lenses. Depending of the nemertean species, from two to several hundreds of eyes can be found in the animal. Some eyeless nemertean species are described. In *Lineus sanguineus* (*L.sanguineus*), the eye number varies from two to more than ten. They are dorsally located and situated beneath the epidermis, in the dermis. They consist of a single-layered epithelium, curved into a cup shape (Fig. 1.8). The cytoplasm of these cells contains pigment granules, brown ones in the case of *L.sanguineus*. The interior of the cup contains photoreceptor cells, which terminate in a rod border that is in direct contact with the pigmented cells. The nerve fibers, from the photoreceptor cells, pass through the mouth of the cup. Hence, similarly to the photoreceptor cells orientation in the vertebrate neural retina, the photoreceptor cells of nemerteans are oriented away from the light signal.

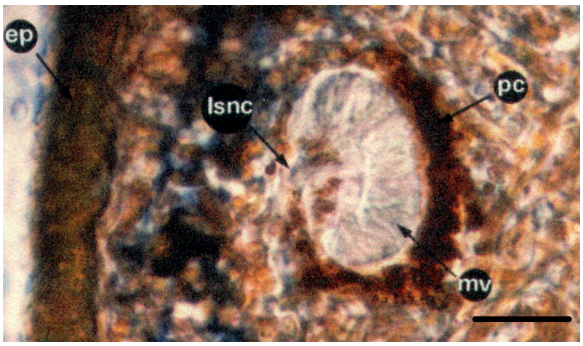


Fig. 1.8 Paraffin section of a *Lineus ruber* eye (From Tarpin *et al.,* 2002)
The eyes of *Lineus* consist of small number of inverted photoreceptor cells into a cup of pigmented cells. ep: epidermis ; pc: pigment cells ; mv: microvilli and lsnc: light sensitive nerve cells. Scale bar : 2,5 μm.

The majority of nemerteans are dioecious (having two separate sexes), but some hermaphrodites have also been described. The process of fertilization, which usually takes place in spring, is mostly external. It can also be internal as in the case of the ovoviviparous terrestrial nemertean, *Geonemertes agricola*. The embryonic development is either direct or indirect, involving a pilidium, an Iwata (found only in *Micrura akkeshiensis*) or a Desor larva. The pilidium and the Iwata larva are similar to the annelid trochophora larva and are free swimming. In contrast, the Desor larva develops and metamorphoses inside the egg membrane. A description of the embryonic development of *L.sanguineus* is not available, probably because this animal reproduces asexually in captivity. *Lineus viridis* (*L.viridis*), a close relative, develops into a Desor larva and subsequently undergoes metamorphosis to gain the adult form. In contrast to the pilidium larva, the Iwata and Desor larva do not feed. Nemerteans share the spiral cleavage type with plathyhelmintes, annelids and mollusks. Few nemertean species show asexual reproduction through fragmentation of their body and subsequent regeneration of the missing structures.

4.2 Regeneration and developmental plasticity in *Lineus*

In higher metazoans there is a progressive loss of the capacity for regeneration of missing body parts in the course of development. However, in a few adult animals the capacity for developmental regulation and regeneration of missing body parts is retained throughout adult life. The best-known examples are Hydra, planaria, cockroaches and salamanders (Trembley, 1744; Spallanzani, 1769; Morgan, 1904; Wolpert *et al.*, 1974; Bryant *et al.*, 1977).

All nemertean species show a certain ability to regenerate damaged body parts. The regeneration capacity depends both on the species and the body region concerned. Most nemerteans can undergo posterior regeneration in order to replace the missing body parts, but very few can regenerate structures anterior to the damaged region. *L.sanguineus* can regenerate all anterior structures, including the complete head and cerebral ganglia. A single adult worm can be cut into several pieces, each regenerating a complete worm. During regeneration, parenchymal phagocytes resorb the old tissues of the worm in order to provide nutrients for the proliferating cells of the regenerative blastema. The presence of the original nerve cord is the only requirement for complete regeneration. In accordance with this dependence, it has been suggested that an organizational center might be present at the level of the nerve cord (Coe, 1932). It has been shown that extracts of head-regenerating blastemata inhibit the anterior regeneration of *Lineus vegetus*, while extracts of posterior blastemata inhibit posterior replacement (Tucker, 1959). It has therefore been proposed that regeneration in *Lineus vegetus* involves a series of differentiation centres, which extend sequentially from the site of the initial regeneration. Tucker has proposed that every organizer center leads to a characteristic differentiation pattern. In addition, she proposed that each center is unable to obtain levels of tissue organization that have already been reached because of inhibitory effects of the centre that has differentiated just before it (Tucker, 1959). Nemertines of the genus *Lineus* not only have a remarkable capacity for regeneration (Dawyddoff, 1942; Gontcharoff, 1951; Bierne, 1962; Bierne, 1970) but also for reconstitution of the normal body pattern by removal of additional body parts by a mechanism called transgeneration (Bierne, 1985; Bierne, 1988; Bierne, 1990; Tarpin *et al.*, 1999; Tarpin *et al.*, 2002).

4.3. *LsHox* and *LsPax-6* homologues

The genome of *L.sanguineus* contains at least six *Hox* genes (Fig. 1.9), which have been assigned to an orthologous group based on their homeobox and on flanking sequences (Fig 1.9): two from the anterior class, homologous to the vertebrate *Hox* 1 and *Hox* 3; three from the middle class, homologous to the vertebrate *Hox* 4, 6, 7 and one from the posterior class, homologous to the vertebrate *Hox* 9 (Kmita-Cunisse *et al.*, 1998).

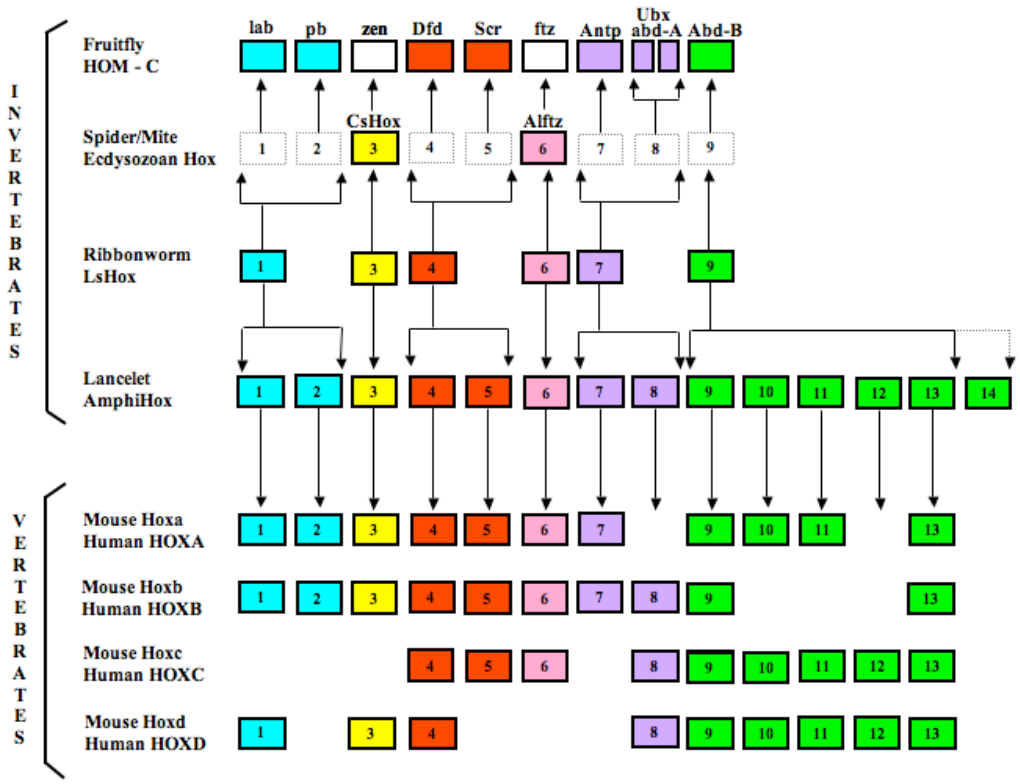


Fig. 1.9 Hypothetical evolution of the *L.sanguineus Hox* cluster
The ribbonworm *L.sanguineus* genome contains at least six *Hox* genes, *LsHox,* which have been assigned to orthologous groups based on comparisons of homeobox and flanking sequences.

The six *L.sanguineus Hox* genes are probably arranged in a single cluster as they have been resolved by pulse-field electrophoresis in a relatively small DNA fragment (Kmita-Cunisse *et al.*, 1998).

A *Pax-6* homologue has also been cloned from *L.sanguineus*: *LsPax-6* (Loosli *et al.,* 1996). *Pax-6* encodes a transcription factor, which contains two evolutionary highly conserved domains, a paired domain and a homeodomain. Among other important functions during development, Pax-6 has been proposed to be a "master control gene" for eye development (Gehring and Ikeo, 1999). *LsPax-6* is expressed in the regenerating brain, cerebral organs and eyes of *L.sanguineus*. *Pax-6* expression is also detected during the development of the brain and the cerebral organs of *L.viridis* (Loosli *et al.*, 1996).

5. Why working with *Lineus*

Throughout this general introduction, I have already mentioned the relevant features of *Lineus* worms that make them models of choice for developmental biology studies. Here, I summarize the three major relevant features of *Lineus*:

Lineus worms are members of the Lophotrochozoa, one of the three major branches of the bilaterian animals. As already mentioned, the "classical" model organisms used for molecular studies belong to only two of the bilaterian branches: the Deuterostomia and the Ecdysozoa. Lophotrochozoan animals have been widely neglected in molecular investigations, although it is widely accepted that most of their development and body plans exhibit evolutionary ancient characteristics (Tessmar-Raible and Arendt, 2003). In agreement with this idea, also *Lineus* has a simple, prototypical body organization along its A-P axis (Fig. 1.7). *Lineus* presents interesting characteristics, such as lateral nerve cords. In addition, many lophotrochozoan genes are more similar at both, intron and exon levels, to vertebrate homologues than to any ecdysozoan ones. Compared to organisms like dipterans, nematodes or ascidians, vertebrates and Lophotrochozoa are considered slow-evolving animals. Thus, it is assumed that the urbilaterian genes were probably more similar in both, structure and sequence, to the genes presently found in vertebrates and Lophotrochozoa than to the one from today's ecdysozoan animals, for example (Raible *et al.*, 2005). In agreement with this notion, all the *L.sanguineus* genes identified so far display closer similarities to the vertebrate homologues than to the ecdysozoan ones (Loosli *et al.,* 1996; Kmita-Cunisse *et al.*, 1998; this PhD work). Hence, studies from Lophotrochozoa might help to reconstruct the urbilaterian genome.

In higher metazoans, there is a progressive loss of the capacity for regeneration of missing body parts in the course of development. However, in the nemertean *L.sanguineus*, the capacity for regeneration and developmental regulation is retained throughout the adult life. Indeed, an adult *L.sanguineus* worm is able to regenerate lost parts of their body and are also able to eliminate grafted surplus tissues (Dawyddoff, 1942; Gontcharoff, 1951; Bierne, 1962; Bierne, 1970). In addition, it shows an impressive developmental plasticity: when it is well fed, its body

continuously grows, whereas upon prolonged periods of starvation, its body shrinks. Furthermore, a *L.sanguineus* worm is capable of developmental reprogramming: when discontinuities are introduced in the positional values along its A-P axis by grafting procedures, cells from differentiated tissues are capable to restore a normal body pattern by different mechanisms, such as intercalary regeneration, transgeneration and transdifferentiation (Bierne, 1985; Bierne, 1988; Bierne, 1990; Tarpin *et al.*, 1999; Tarpin *et al.*, 2002). Because of its relative simplicity, developmental plasticity and its evolutionary position, *L.sanguineus* is an attractive system to investigate the specification of the body plan and the mechanism by which differentiated cells maintain or reprogram their identity in a context-dependent manner. Regeneration, like development, involves the self-assembly of new tissues. But, in contrast to development, regeneration requires an anatomical and functional integration of the newly formed body parts into the pre-existing tissues. Hence, studying regeneration might not only bring light on topics, such as tissue polarity, patterning but also on topics, such as the control of size and proportions and the establishment of connections between adult CNS parts and juvenile, regenerated CNS ones.

In contrast to *L.sanguineus*, many species, especially higher vertebrates, do not respond to injury or tissue removal by re-growing missing body parts. We need to understand why the regeneration capacity is lost in these animals in order to know whether, in the future, a regenerative process could be induced, at least to some extent, in these animals. A good approach to answer this question is to molecularly compare close species that do not exhibit the same regeneration capacities. It is noteworthy that members of the same nemertean genus *Lineus* respond very differently to amputation: for example, while a *L.sanguineus* can regenerate its brain, *L.lacteus* cannot. By comparing the molecular events that follow an amputation in nemerteans capable of regeneration with the ones occurring in nemerteans incapable of regeneration, we hope to find some differences that could explain the variation in regeneration capacities of these two species.

6. References of the general introduction

Aboobaker, A. A. and Blaxter, M. L. (2003). "*Hox* Gene Loss during Dynamic Evolution of the Nematode Cluster." *Curr Biol* **13**(1): 37-40.

Abzhanov, A. and Kaufman, T. C. (2000). "Crustacean (malacostracan) *Hox* genes and the evolution of the arthropod trunk." *Development* **127**(11): 2239-2249.

Adoutte, A., Balavoine, G., Lartillot, N., Lespinet, O., Prud'homme, B. and de Rosa, R. (2000). "The new animal phylogeny: reliability and implications." *Proc Natl Acad Sci U S A* **97**(9): 4453-4456.

Affolter, M., Percival-Smith, A., Muller, M., Leupin, W. and Gehring, W. J. (1990). "DNA binding properties of the purified Antennapedia homeodomain." *Proc Natl Acad Sci U S A* **87**(11): 4093-4097.

Aguinaldo, A. M., Turbeville, J. M., Linford, L. S., Rivera, M. C., Garey, J. R., Raff, R. A. and Lake, J. A. (1997). "Evidence for a clade of nematodes, arthropods and other moulting animals." *Nature* **387**(6632): 489-493.

Akam, M. (1998). "*Hox* genes: from master genes to micromanagers." *Curr Biol* **8**(19): R676-678.

Averof, M. and Patel, N. H. (1997). "Crustacean appendage evolution associated with changes in *Hox* gene expression." *Nature* **388**(6643): 682-686.

Bierne, J. (1962). "La régénération de la trompe chez les Nemertes." *Bull. Biol. Fr. Belg.* **96**: 481-504.

Bierne, J. (1970). "Recherches sur la différenciation sexuelle au cours de l'ontogenèse et de la régénération chez le némertien *Lineus ruber*." *Ann. Sc. Nat. Zool.* **12**: 181-298.

Bierne, J. (1985). "Histocompatibility in nemertines: fates of multiparental *Lineus* constructed by grafting of pieces from many donors." *Am. Zool.* **125**: 135-144.

Bierne, J. (1988). Transgeneration as an alternative to regeneration in heterotopically reconstructed *Lineus.* Control of cell proliferation and differentiation during regeneration. H. J. K. Anton, Basel)**:** p 230-234.

Bierne, J. (1990). "*Lineus* as a model for studying developmental processes in animals reconstructed from adult pieces." *Int J Dev Biol* **34**(1): 245-253.

Biggin, M. D. and McGinnis, W. (1997). "Regulation of segmentation and segmental identity by *Drosophila* homeoproteins: the role of DNA binding in functional activity and specificity." *Development* **124**(22): 4425-4433.

Billeter, M., Qian, Y. Q., Otting, G., Muller, M., Gehring, W. and Wuthrich, K. (1993). "Determination of the nuclear magnetic resonance solution structure of an Antennapedia homeodomain-DNA complex." *J Mol Biol* **234**(4): 1084-1093.

Bridges, C. B. and Morgan, T. H. (1923). The third-chromosome group of mutant characters of *Drosophila melanogaster*. C. I. o. Washington**:** pp. 137-146.

Brooke, N. M., Garcia-Fernandez, J. and Holland, P. W. (1998). "The *ParaHox* gene cluster is an evolutionary sister of the *Hox* gene cluster." *Nature* **392**(6679): 920-922.

Brunet, I., Weinl, C., Piper, M., Trembleau, A., Volovitch, M., Harris, W., Prochiantz, A. and Holt, C. (2005). "The transcription factor Engrailed-2 guides retinal axons." *Nature* **438**(7064): 94-98.

Brusca, R. and Brusca, G. (1990). Invertebrates. Sinauer, Sunderland, Mass.

Bryant, P. J., Bryant, S. V. and French, V. (1977). "Biological regeneration and pattern formation." *Sci Am* **237**: 66-81.

Burglin, T. R. and Ruvkun, G. (1993). "The *Caenorhabditis elegans* homeobox gene cluster." *Curr Opin Genet Dev* **3**(4): 615-620.

Burke, A. C., Nelson, C. E., Morgan, B. A. and Tabin, C. (1995). "*Hox* genes and the evolution of vertebrate axial morphology." *Development* **121**(2): 333-346.

Carranza, S., Baguna, J. and Riutort, M. (1997). "Are the Platyhelminthes a monophyletic primitive group? An assessment using 18S rDNA sequences." *Mol Biol Evol* **14**(5): 485-497.

Carrasco, A. E., McGinnis, W., Gehring, W. J. and De Robertis, E. M. (1984). "Cloning of an *X. laevis* gene expressed during early embryogenesis coding for a peptide region homologous to *Drosophila* homeotic genes." *Cell* **37**(2): 409-414.

Chambeyron, S. and Bickmore, W. A. (2004). "Chromatin decondensation and nuclear reorganization of the *HoxB* locus upon induction of transcription." *Genes Dev* **18**(10): 1119-1130.

Chourrout, D., Delsuc, F., Chourrout, P., Edvardsen, R. B., Rentzsch, F., Renfer, E., Jensen, M. F., Zhu, B., de Jong, P., Steele, R. E. and Technau, U. (2006). "Minimal *ProtoHox* cluster inferred from bilaterian and cnidarian *Hox* complements." *Nature* **442**(7103): 684-687.

Coe, W. R. (1932). "Regeneration in nemerteans. III Regeneration of *Lineus pictifrons*." *J. Exp. Zool* **61**: 29-43.

Cook, C. E., Jimenez, E., Akam, M. and Salo, E. (2004). "The *Hox* gene complement of acoel flatworms, a basal bilaterian clade." *Evol Dev* **6**(3): 154-163.

Crow, K. D., Stadler, P. F., Lynch, V. J., Amemiya, C. and Wagner, G. P. (2006). "The "fish-specific" *Hox* cluster duplication is coincident with the origin of teleosts." *Mol Biol Evol* **23**(1): 121-136.

Darling, J. A., Reitzel, A. R., Burton, P. M., Mazza, M. E., Ryan, J. F., Sullivan, J. C. and Finnerty, J. R. (2005). "Rising starlet: the starlet sea anemone, *Nematostella vectensis*." *Bioessays* **27**(2): 211-221.

Dawyddoff, C. (1942). "Régénération créatrice chez les Némertes." *Bull. Biol. Fr. Belg.* **76**: 59-191.

De Robertis, E. M. and Sasai, Y. (1996). "A common plan for dorsoventral patterning in Bilateria." *Nature* **380**(6569): 37-40.

De Robertis, E. M. (1997). "Evolutionary biology. The ancestry of segmentation." *Nature* **387**(6628): 25-26.

De Rosa, R., Grenier, J. K., Andreeva, T., Cook, C. E., Adoutte, A., Akam, M., Carroll, S. B. and Balavoine, G. (1999). "*Hox* genes in brachiopods and priapulids and protostome evolution." *Nature* **399**(6738): 772-776.

Duboule, D. (1991). "Patterning in the vertebrate limb." *Curr Opin Genet Dev* **1**(2): 211-216.

Duboule, D. (1992). "The vertebrate limb: a model system to study the *Hox/HOM* gene network during development and evolution." *Bioessays* **14**(6): 375-384.

Duboule, D. (1994). "Temporal colinearity and the phylotypic progression: a basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony." *Dev Suppl*: 135-142.

Duboule, D. and Morata, G. (1994). "Colinearity and functional hierarchy among genes of the homeotic complexes." *Trends Genet* **10**(10): 358-364.

Ferrier, D. E. and Holland, P. W. (2001). "Ancient origin of the *Hox* gene cluster." *Nat Rev Genet* **2**(1): 33-38.

Ferrier, D. E. and Minguillon, C. (2003). "Evolution of the *Hox/ParaHox* gene clusters." *Int J Dev Biol* **47**(7-8): 605-611.

Finnerty, J. R., Pang, K., Burton, P., Paulson, D. and Martindale, M. Q. (2004). "Origins of bilateral symmetry: *Hox* and *dpp* expression in a sea anemone." *Science* **304**(5675): 1335-1337.

Garber, R. L, Kuroiwa, A., Gehring, W. J. (1983). "Genomic and cDNA clones of the homeotic locus *Antennapedia* in *Drosophila.*" *Embo J*. **2**(11): 2027-2036.

Garcia-Fernandez, J. and Holland, P. W. (1994). "Archetypal organization of the amphioxus *Hox* gene cluster." *Nature* **370**(6490): 563-566.

Garcia-Fernandez, J. (2005a). "*Hox, ParaHox*, *ProtoHox*: facts and guesses." *Heredity* **94**(2): 145-152.

Garcia-Fernandez, J. (2005b). "The genesis and evolution of homeobox gene clusters." *Nat Rev Genet* **6**(12): 881-892.

Gauchat, D., Mazet, F., Berney, C., Schummer, M., Kreger, S., Pawlowski, J. and Galliot, B. (2000). "Evolution of *Antp*-class genes and differential expression of Hydra *Hox/paraHox* genes in anterior patterning." *Proc Natl Acad Sci U S A* **97**(9): 4493-4498.

Gehring, W. (1966). "Bildung eines vollständigen Mittelbeines mit Sternopleura in der antennenregion bei der Mutante Nasobemia (Ns) von *Drosophila melanogaster*." *Jul. Klaus Arch.* **41**: 44-54.

Gehring, W. J. (1993). "Exploring the homeobox." *Gene* **135**(1-2): 215-221.

Gehring, W. J., Affolter, M. and Burglin, T. (1994a). "Homeodomain proteins." *Annu Rev Biochem* **63**: 487-526.

Gehring, W. J., Qian, Y. Q., Billeter, M., Furukubo-Tokunaga, K., Schier, A. F., Resendez-Perez, D., Affolter, M., Otting, G. and Wuthrich, K. (1994b). "Homeodomain-DNA recognition." *Cell* **78**(2): 211-223.

Gehring, W. J. (1998). Master Control Genes in Development and Evolution: The Homeobox Story. Yale University Press. New Haven, USA.

Gehring, W. J. and Ikeo, K. (1999). "Pax 6: mastering eye morphogenesis and eye evolution." *Trends Genet* **15**(9): 371-377.

Gellon, G. and McGinnis, W. (1998). "Shaping animal body plans in development and evolution by modulation of *Hox* expression patterns." *Bioessays* **20**(2): 116-125.

Gibson, R., Ed. (1972). Nemerteans, *Hutchinson Publ. London*.

Gilbert, S. F. (2003). "Opening Darwin's black box: teaching evolution through developmental genetics." *Nat Rev Genet* **4**(9): 735-741.

Gontcharoff, M. (1951). "Biologie de la régénération et de la reproduction chez quelques *Lineidae* de France." *Ann. Soc. Nat. Zool.* **13**: 149-235.

Gonzalez-Reyes, A., Urquia, N., Gehring, W. J., Struhl, G. and Morata, G. (1990). "Are cross-regulatory interactions between homoeotic genes functionally significant?" *Nature* **344**(6261): 78-80.

Halanych, K. M., Bacheller, J. D., Aguinaldo, A. M., Liva, S. M., Hillis, D. M. and Lake, J. A. (1995). "Evidence from 18S ribosomal DNA that the lophophorates are protostome animals." *Science* **267**(5204): 1641-1643.

Holland, P. W., Garcia-Fernandez, J., Williams, N. A. and Sidow, A. (1994). "Gene duplications and the origins of vertebrate development." *Dev Suppl*: 125-133.

Holland, P. W. and Takahashi, T. (2005). "The evolution of homeobox genes: Implications for the study of brain development." *Brain Res Bull* **66**(4-6): 484-490.

Hombria, J. C. and Lovegrove, B. (2003). "Beyond homeosis--HOX function in morphogenesis and organogenesis." *Differentiation* **71**(8): 461-476.

Hurst, L. D., Pal, C. and Lercher, M. J. (2004). "The evolutionary dynamics of eukaryotic gene order." *Nat Rev Genet* **5**(4): 299-310.

Ikuta, T., Yoshida, N., Satoh, N. and Saiga, H. (2004). "Ciona intestinalis *Hox* gene cluster: Its dispersed structure and residual colinear expression in development." *Proc Natl Acad Sci U S A* **101**(42): 15118-15123.

Iwata, F. (1985). "Foregut formation of the Nemerteans and its role in Nemertean systematics" *Amer. Zool.* **25**: 23–36.

Jagla, K., Bellard, M. and Frasch, M. (2001). "A cluster of *Drosophila* homeobox genes involved in mesoderm differentiation programs." *Bioessays* **23**(2): 125-133.

Joliot, A., Maizel, A., Rosenberg, D., Trembleau, A., Dupas, S., Volovitch, M. and Prochiantz, A. (1998). "Identification of a signal sequence necessary for the unconventional secretion of Engrailed homeoprotein." *Curr Biol* **8**(15): 856-863.

Kamm, K., Schierwater, B., Jakob, W., Dellaporta, S. L. and Miller, D. J. (2006). "Axial patterning and diversification in the cnidaria predate the *Hox* system." *Curr Biol* **16**(9): 920-926.

Kaufman, T. C., Seeger, M. A. and Olsen, G. (1990). "Molecular and genetic organization of the *antennapedia* gene complex of *Drosophila melanogaster*." *Adv Genet* **27**: 309-362.

Kessel, M. and Gruss, P. (1991). "Homeotic transformations of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid." *Cell* **67**(1): 89-104.

Kessel, M. (1992). "Respecification of vertebral identities by retinoic acid." *Development* **115**(2): 487-501.

Keys, D. N., Lewis, D. L., Selegue, J. E., Pearson, B. J., Goodrich, L. V., Johnson, R. L., Gates, J., Scott, M. P. and Carroll, S. B. (1999). "Recruitment of a hedgehog regulatory circuit in butterfly eyespot evolution." *Science* **283**(5401): 532-534.

Kissinger, C. R., Liu, B. S., Martin-Blanco, E., Kornberg, T. B. and Pabo, C. O. (1990). "Crystal structure of an engrailed homeodomain-DNA complex at 2.8 A resolution: a framework for understanding homeodomain-DNA interactions." *Cell* **63**(3): 579-590.

Kmita-Cunisse, M., Loosli, F., Bierne, J. and Gehring, W. J. (1998). "Homeobox genes in the ribbonworm *Lineus sanguineus*: evolutionary implications." *Proc Natl Acad Sci U S A* **95**(6): 3030-3035.

Kmita, M. and Duboule, D. (2003). "Organizing axes in time and space; 25 years of colinear tinkering." *Science* **301**(5631): 331-333.

Kortschak, R. D., Samuel, G., Saint, R. and Miller, D. J. (2003). "EST analysis of the cnidarian *Acropora millepora* reveals extensive gene loss and rapid sequence divergence in the model invertebrates." *Curr Biol* **13**(24): 2190-2195.

Kusserow, A., Pang, K., Sturm, C., Hrouda, M., Lentfer, J., Schmidt, H. A., Technau, U., von Haeseler, A., Hobmayer, B., Martindale, M. Q. and Holstein, T. W. (2005). "Unexpected complexity of the *Wnt* gene family in a sea anemone." *Nature* **433**(7022): 156-160.

Lee, P. N., Callaerts, P., De Couet, H. G. and Martindale, M. Q. (2003). "Cephalopod *Hox* genes and the origin of morphological novelties." *Nature* **424**(6952): 1061-1065.

Levine, M., Rubin, G. M. and Tjian, R. (1984). "Human DNA sequences homologous to a protein coding region conserved between homeotic genes of *Drosophila*." *Cell* **38**(3): 667-673.

Lewis, E. B. (1978). "A gene complex controlling segmentation in *Drosophila*." *Nature* **276**(5688): 565-570.

Li, X., Murre, C. and McGinnis, W. (1999). "Activity regulation of a Hox protein and a role for the homeodomain in inhibiting transcriptional activation." *Embo J* **18**(1): 198-211.

Loosli, F., Kmita-Cunisse, M. and Gehring, W. J. (1996). "Isolation of a *Pax-6* homolog from the ribbonworm *Lineus sanguineus*." *Proc Natl Acad Sci U S A* **93**(7): 2658-2663.

Lowe, C. J. and Wray, G. A. (1997). "Radical alterations in the roles of homeobox genes during echinoderm evolution." *Nature* **389**(6652): 718-721.

Luke, G. N., Castro, L. F., McLay, K., Bird, C., Coulson, A. and Holland, P. W. (2003). "Dispersal of *NK* homeobox gene clusters in amphioxus and humans." *Proc Natl Acad Sci U S A* **100**(9): 5292-5295.

Mann, R. S. and Morata, G. (2000). "The developmental and molecular biology of genes that subdivide the body of *Drosophila*." *Annu Rev Cell Dev Biol* **16**: 243-271.

Manuel, M., Jager, M., Murienne, J., Clabaut, C. and Guyader, H. L. (2006). "*Hox* genes in sea spiders (*Pycnogonida*) and the homology of arthropod head segments." *Dev Genes Evol* **216**(7-8): 481-491.

McGinnis, W., Levine, M. S., Hafen, E., Kuroiwa, A. and Gehring, W. J. (1984a). "A conserved DNA sequence in homeotic genes of the *Drosophila Antennapedia* and *bithorax* complexes." *Nature* **308**(5958): 428-433.

McGinnis, W., Garber, R. L., Wirz, J., Kuroiwa, A. and Gehring, W. J. (1984b). "A homologous protein-coding sequence in *Drosophila* homeotic genes and its conservation in other metazoans." *Cell* **37**(2): 403-408.

McGinnis, W. and Krumlauf, R. (1992). "Homeobox genes and axial patterning." *Cell* **68**(2): 283-302.

Merabet, S., Kambris, Z., Capovilla, M., Berenger, H., Pradel, J. and Graba, Y. (2003). "The hexapeptide and linker regions of the AbdA Hox protein regulate its activating and repressive functions." *Dev Cell* **4**(5): 761-768.

Merabet, S., Pradel, J. and Graba, Y. (2005). "Getting a molecular grasp on Hox contextual activity." *Trends Genet* **21**(9): 477-480.

Moens, C. B. and Selleri, L. (2006). "Hox cofactors in vertebrate development." *Dev Biol* **291**(2): 193-206.

Monteiro, A. S. and Ferrier, D. E. (2006). "*Hox* genes are not always Colinear." *Int J Biol Sci* **2**(3): 95-103.

Morgan, T. H. (1904). "Regeneration of heteromorphic tails in posterior pieces of *Planaria simplicissima.*" *J Exp Zool 1*: 385-394.

Nedelec, S., Foucher, I., Brunet, I., Bouillot, C., Prochiantz, A. and Trembleau, A. (2004). "Emx2 homeodomain transcription factor interacts with eukaryotic translation initiation factor 4E (eIF4E) in the axons of olfactory sensory neurons." *Proc Natl Acad Sci U S A* **101**(29): 10815-10820.

Niessing, D., Driever, W., Sprenger, F., Taubert, H., Jackle, H. and Rivera-Pomar, R. (2000). "Homeodomain position 54 specifies transcriptional versus translational control by Bicoid." *Mol Cell* **5**(2): 395-401.

Nusbaum, J. and Oxner, M. (1913). " Die embryonalentwicklung des *Lineus* ruber (Müller)." Z. *Wiss. Zool.* **107**: 78–194.

Otting, G., Qian, Y. Q., Billeter, M., Muller, M., Affolter, M., Gehring, W. J. and Wuthrich, K. (1990). "Protein--DNA contacts in the structure of a homeodomain--DNA complex determined by nuclear magnetic resonance spectroscopy in solution." *Embo J* **9**(10): 3085-3092.

Patel, N. H. (2004). "Evolutionary biology: time, space and genomes." *Nature* **431**(7004): 28-29.

Philippe, H., Lartillot, N. and Brinkmann, H. (2005). "Multigene analyses of bilaterian animals corroborate the monophyly of Ecdysozoa, Lophotrochozoa, and Protostomia." *Mol Biol Evol* **22**(5): 1246-1253.

Plaza, S., Prince, F., Jaeger, J., Kloter, U., Flister, S., Benassayag, C., Cribbs, D. and Gehring, W. J. (2001). "Molecular basis for the inhibition of *Drosophila* eye development by Antennapedia." *Embo J* **20**(4): 802-811.

Prochiantz, A. and Joliot, A. (2003). "Can transcription factors function as cell-cell signalling molecules?" *Nat Rev Mol Cell Biol* **4**(10): 814-819.

Qian, Y. Q., Billeter, M., Otting, G., Muller, M., Gehring, W. J. and Wuthrich, K. (1989). "The structure of the Antennapedia homeodomain determined by NMR spectroscopy in solution: comparison with prokaryotic repressors." *Cell* **59**(3): 573-580.

Raible, F. and Arendt, D. (2004). "Metazoan evolution: some animals are more equal than others." *Curr Biol* **14**(3): R106-108.

Raible, F., Tessmar-Raible, K., Osoegawa, K., Wincker, P., Jubin, C., Balavoine, G., Ferrier, D., Benes, V., de Jong, P., Weissenbach, J., Bork, P. and Arendt, D. (2005). "Vertebrate-type intron-rich genes in the marine annelid *Platynereis dumerilii*." *Science* **310**(5752): 1325-1326.

Rivera-Pomar R, Niessing D, Schmidt-Ott U, Gehring W. J, and Jackle H. (1996). "RNA binding and translational suppression by bicoid." *Nature* **379**(6567): 746-749.

Sanchez Alvarado, A. (2000). "Regeneration in the metazoans: why does it happen?" *Bioessays* **22**(6): 578-590.

Sanchez-Herrero, E., Vernos, I., Marco, R. and Morata, G. (1985). "Genetic organization of *Drosophila bithorax* complex." *Nature* **313**(5998): 108-113.

Schier, A. F. and Gehring, W. J. (1992). "Direct homeodomain-DNA interaction in the autoregulation of the *fushi tarazu* gene." *Nature* **356**(6372): 804-807.

Schock, F., Reischl, J., Wimmer, E., Taubert, H., Purnell, B. A. and Jackle, H. (2000). "Phenotypic suppression of empty spiracles is prevented by buttonhead." *Nature* **405**(6784): 351-354.

Scott, M. P. and Weiner, A. J., Hazelrigg, T. I., Polisky, B. A., Pirrotta, V., Scalenghe, F. and Kaufman, T. C. (1983). "The molecular organization of the *Antennapedia* locus of *Drosophila*." *Cell* **35**: 763-776.

Scott, M. P. and Weiner, A. J. (1984). "Structural relationships among genes that control development: sequence homology between the *Antennapedia*, *Ultrabithorax*, and *fushi tarazu* loci of *Drosophila*." *Proc Natl Acad Sci U S A* **81**(13): 4115-4119.

Seimiya, M., Ishiguro, H., Miura, K., Watanabe, Y. and Kurosawa, Y. (1994). "Homeobox-containing genes in the most primitive metazoa, the sponges." *Eur J Biochem* **221**(1): 219-225.

Seo, H. C., Edvardsen, R. B., Maeland, A. D., Bjordal, M., Jensen, M. F., Hansen, A., Flaat, M., Weissenbach, J., Lehrach, H., Wincker, P., Reinhardt, R. and Chourrout, D. (2004). "*Hox* cluster disintegration with persistent anteroposterior order of expression in *Oikopleura dioica*." *Nature* **431**(7004): 67-71.

Shepherd, J. C., McGinnis, W., Carrasco, A. E., De Robertis, E. M. and Gehring, W. J. (1984). "Fly and frog homeodomains show homologies with yeast mating type regulatory proteins." *Nature* **310**(5972): 70-71.

Slack, J. M., Holland, P. W. and Graham, C. F. (1993). "The zootype and the phylotypic stage." *Nature* **361**(6412): 490-492.

Sommer, H., Beltran, J. P., Huijser, P., Pape, H., Lonnig, W. E., Saedler, H. and Schwarz-Sommer, Z. (1990). "*Deficiens*, a homeotic gene involved in the control of flower morphogenesis in *Antirrhinum majus*: the protein shows homology to transcription factors." *Embo J* **9**(3): 605-613.

Spallanzani, L. (1769). Prodromo di un opera da imprimersi sopra la riproduzioni animali (An essay on animal reproduction).

Stern, C. D., Charite, J., Deschamps, J., Duboule, D., Durston, A. J., Kmita, M., Nicolas, J. F., Palmeirim, I., Smith, J. C. and Wolpert, L. (2006). "Head-tail patterning of the vertebrate embryo: one, two or many unresolved problems?" *Int J Dev Biol* **50**(1): 3-15.

Tarchini, B. and Duboule, D. (2006). "Control of *Hoxd* genes' collinearity during early limb development." *Dev Cell* **10**(1): 93-103.

Tarpin, M., Gehring, W. J. and Bierne, J. (1999). "Reverse homeosis in homeotically reconstructed ribbonworms." *Proc Natl Acad Sci U S A* **96**(21): 11900-11903.

Tarpin, M., Gehring, W. J. and Bierne, J. (2002). "Conversion of a postocellar into an ocellar region as a transdetermination event occurring in adult ribbonworms." *Mech Dev* **118**(1-2): 39-44.

Tessmar-Raible, K. and Arendt, D. (2003). "Emerging systems: between vertebrates and arthropods, the Lophotrochozoa." *Curr Opin Genet Dev* **13**(4): 331-340.

Topisirovic, I., Culjkovic, B., Cohen, N., Perez, J. M., Skrabanek, L. and Borden, K. L. (2003). "The proline-rich homeodomain protein, PRH, is a tissue-specific inhibitor of eIF4E-dependent cyclin D1 mRNA transport and growth." *Embo J* **22**(3): 689-703.

Treisman, J., Gonczy, P., Vashishtha, M., Harris, E. and Desplan, C. (1989). "A single amino acid can determine the DNA binding specificity of homeodomain proteins." *Cell* **59**(3): 553-562.

Trembley, A. (1744). Mémoires pour servir à l'histoire d'un genre de polypes d'eau douce.

Tucker, M. (1959). "Inhibitory control of regeneration in nemertean worms." *J. Morph.* **105**: 569-599.

Turbeville, J. M. (1991). Nemertinea. Platyhelminthes and Nemertinea, in: Microscopic anatomy of invertebrates eds. F.W. Harrison and B.J. Bogitsh pp: 285-328

Turbeville, J. M., Field, K. G. and Raff, R. A. (1992). "Phylogenetic position of phylum Nemertini, inferred from 18S rRNA sequences: molecular data as a test of morphological character homology." *Mol Biol Evol* **9**(2): 235-249.

Wagner, G. P., Amemiya, C. and Ruddle, F. (2003). "*Hox* cluster duplications and the opportunity for evolutionary novelties." *Proc Natl Acad Sci U S A* **100**(25): 14603-14606.

Williams, T. M., Williams, M. E. and Innis, J. W. (2005). "Range of HOX/TALE superclass associations and protein domain requirements for HOXA13:MEIS interaction." *Dev Biol* **277**(2): 457-471.

Williams, M. E., Lehoczky, J. A. and Innis, J. W. (2006). "A group 13 homeodomain is neither necessary nor sufficient for posterior prevalence in the mouse limb." *Dev Biol*. **297**(2): 493-507.

Wolberger, C., Vershon, A. K., Liu, B., Johnson, A. D. and Pabo, C. O. (1991). "Crystal structure of a MAT alpha 2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions." *Cell* **67**(3): 517-528.

Wolpert, L., Hornbruchn, A. and Clarke, M. (1974). "Positional information and positional signalling in Hydra." *Amer Zool* **14**: 647-663.

Yanze, N., Spring, J., Schmidli, C. and Schmid, V. (2001). "Conservation of *Hox/ParaHox*-related genes in the early development of a cnidarian." *Dev Biol* **236**(1): 89-98.

Yaron, Y., McAdara, J. K., Lynch, M., Hughes, E. and Gasson, J. C. (2001). "Identification of novel functional regions important for the activity of HOXB7 in mammalian cells." *J Immunol* **166**(8): 5058-5067.

Zhang, H., Catron, K. M. and Abate-Shen, C. (1996). "A role for the Msx homeodomain in transcriptional regulation: residues in the N-terminal arm mediate TATA binding protein interaction and transcriptional repression." *Proc Natl Acad Sci U S A* **93**(5): 1764-1769.

CHAPTER II

Involvement of *Ls-Otx* and *Ls-Cdx,* two homeobox-containing genes, in the antero-posterior patterning of the body axis during both, development and regeneration of *L.sanguineus*

1. Introduction

The combined action of multiple genes is required for the proper specification of axial positions along the embryos. This process begins early in embryogenesis, during gastrulation. Prospective candidates for genes working in the specification of the A-P axis have been identified by molecular and mutational studies performed in various animals. The vertebrate *Hox* genes, like their *Drosophila* homologues, have been found to play a pivotal role in the A-P axis specification (McGinnis and Krumlauf, 1992). Along the A-P axis, the most rostral *Hox* gene expression is posterior to the midbrain-hindbrain boundary. This finding suggests that other genes regulate the development of the anterior head structures. *Otx*, a homeobox-containing gene, is expressed in rostral brain regions in many bilaterian animals and is thought to specify the anterior regions of their embryo (Bally-Cuif and Boncinelli, 1997). *Cdx*, another homeobox-containing non-*Hox* gene, has been suggested to be part of another evolutionary conserved patterning system. In contrast to *Otx*, *Cdx* members preferentially localize to the posterior of the developing embryo (Marom *et al.*, 1997). They have been suggested to play important roles in the specification of the most posterior region of the embryo during gastrulation and neurulation. They probably achieve these functions in part by regulating some *Hox* family members (Pownall *et al.*, 1996).

1.1 The *Otx*/ *Otd* related genes

Conserved regulatory genes, which are commonly expressed in the head and the brain of diverse animals, have been identified in recent molecular studies. They encode for transcription factor that first operate in the early anterior patterning events of the embryo. Secondly, many of them also set up the formation of the brain primordia and the regionalization of the developing brain. Among the anterior patterning genes, the *Otx* genes, encoding paired-class homeodomain proteins, are evolutionary highly conserved. They have been isolated from cnidaria (Muller *et al.*, 1999; Smith *et al.*, 1999) to annelid worms (Bruce and Shankland, 1998) and insects (Finkelstein *et al.*, 1990a), to Amphioxus (Williams and Holland, 1998) and vertebrates, including humans, where three family members, Otx1, Otx2 (Simeone *et al.*, 1992) and Crx (Furukawa *et al.*, 1997), have been identified.

1.1.1 The *Drosophila Otd* gene

The *Drosophila Otx* homologue, *orthodenticle* (*otd*), is first expressed at the anterior pole of the blastoderm embryo, in a broad circumferential stripe, including the antennal and the preantennal procephalic regions of the head. Later, during gastrulation, its expression becomes restricted to the procephalic neurectoderm, where it is expressed in most delaminating neuroblasts of the presumptive protocerebrum and the presumptive anterior deutocerebrum (Cohen and Jürgens, 1990; Finkelstein and Perrimon, 1990b). Interestingly, *otd* is not expressed in the most anterior part of the brain. *otd* expression is also observed along the midline of the developing ventral nerve cord (Wieschaus *et al.*, 1992). In *otd* mutant flies, the protocerebral anlage is completely deleted and some deuterocerebral neuroblasts do not form. Therefore the brain is severely reduced (Hirth *et al.*, 1995; Younossi-Hartenstein *et al.*, 1997). Other abnormalities are also observed in the ventral nerve cord and in non-neural structures. *otd* determines cell fate during both embryonic development, where it is required for the head and ventral midline development, and larval development. Indeed, during larval stages, *otd* is expressed in specific imaginal discs and is required to specify medial cell fates in both the larval and adult epidermis (Wieschaus *et al.*, 1992). In particular, *otd* is required for the development of the dorsal region of the adult head capsule that is formed by the fusion of the two eye-antennal discs. It has been shown that the proper establishment of the three regions of the dorsal head depends on different concentration of OTD protein. Indeed, specific threshold levels of *otd* expression are needed to set up the different subdomains along the mediolateral axis of the adult head (Royet and Finkelstein, 1995).

1.1.2 The mouse *Otx* genes

The two mouse *Otx* homologues, *Otx1* and *Otx2,* are required in early specification and patterning of the anterior neurectoderm (Simeone *et al.*, 1992). They are expressed in nested domains in the forebrain and the midbrain. It is noteworthy that, similar to *otd* expression in the fly embryonic brain, *Otx1* and *Otx2* expression domains in mouse do not include the most anterior brain region (Simeone *et al.*, 1992). The mouse *Otx* homologues also play important roles in neuroblast proliferation, corticogenesis as well as in olfactory, visual and acoustic sense organ

development (Simeone *et al.*, 1993). There are evidences from studies addressing the formation of specific subdomains of the fly head that the mechanism of brain patterning is dependent on appropriate OTX protein concentration levels. Indeed, precise threshold of the mouse Otx1 and Otx2 are required for the distinction of adjacent territories with different fates within the rostral neural tube (Acampora *et al.*, 1997). *Otx1* mutant mice are viable but exhibit defects in the dorsal telencephalic cortex and in the development of the acoustic and visual sense organs. They also suffer from spontaneous epileptic seizures (Acampora *et al.*, 1996). *Otx2* mutant mice display major abnormalities in their body plan and die early in development. As they fail to gastrulate properly, the specification of their anterior neurectoderm does not occur effectively. As a consequence, the forebrain, midbrain and rostral hindbrain regions are completely absent in *Otx2* mutant mice (Acampora *et al.*, 1995; Ang *et al.*, 1996).

2.1.1.3 Otx-related gain-of-funtion

Gain-of-function studies in several animals have confirmed the conserved crucial role of *Otx* in the specification of the embryonic A-P body axis. *Xenopus* embryos microinjected with *Xotx2* mRNA exhibit reductions of the trunk and the tail, a partial secondary axis and ectopic anterior structures, such as cement glands and neural tissue (Blitz and Cho, 1995; Pannese *et al.*, 1995). Ubiquitous overexpression of *otd* in *Drosophila* results in formation of additional anterior neural structures, duplication of anterior sensory structures and suppression of trunk nerve cord development (Gallitano-Mendel and Finkelstein, 1998; Leuzinger *et al.*, 1998). Injection of *Hroth* mRNA into fertilized ascidians eggs results in an expansion of the trunk and in a reduction of the tail structures, via suppression of specific gene expression (Wada and Saiga, 1999).

1.1.4 *Otx* gene cross-phylum rescue experiment

Finding similar *Otx* expression patterns in homologous primordia across different phyla and comparable *Otx* mutant phenotypes argues for conserved functions of this gene among Bilateria. *Otx* plays important roles in the specification of the embryonic axis and in the formation of the brain and other anterior structures. It

also possesses a conserved function in the repression of posterior structures. To corroborate the idea that *Otx*/ *otd* genes have conserved genetic functions, several cross-phylum gene replacement experiments have been carried out (Fig. 2.1).

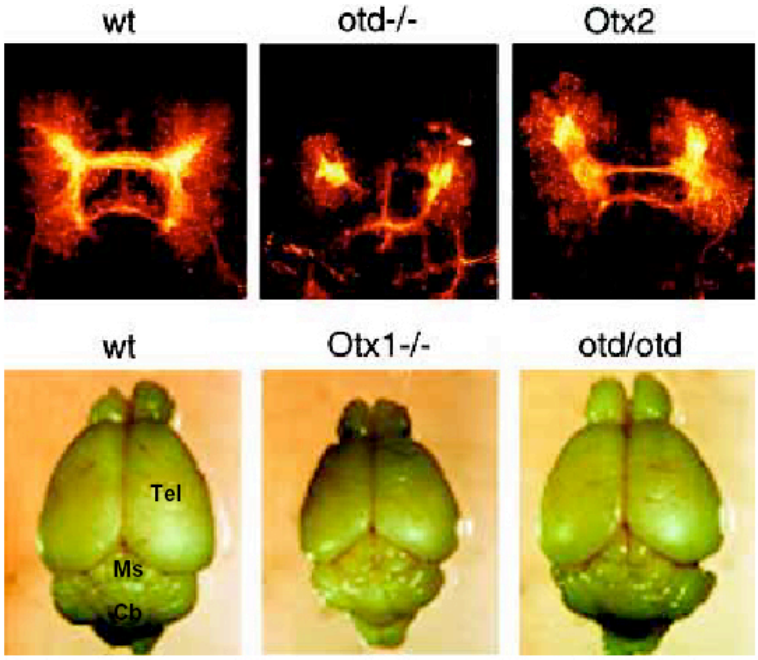


Fig. 2.1 *otd*/ *Otx* cross-phylum rescue experiments in *Drosophila* and the mouse (Reichert, 2002)
Upper panel: In *Drosophila*, the prominent anterior lobes, which are interconnected by an anterior brain commissure, are lost in *otd* null mutant brains. These structures can be restored by overexpressing the human *Otx2* gene in *otd* mutant flies.
Lower panel: In the mouse, the size of the adult *Otx1* null mutant brain is markedly reduced in comparison to a wild type brain. The adult brain size is largely restored by a gene replacement with the *Drosophila otd* homolog.
WT: wild type; Tel: telencephalon; Ms: mesencephalon; Cb: cerebellum.

Heat-shock inductions of the human orthologues *hOtx1* and *hOtx2* in *Drosophila otd* null mutants rescue the CNS defects at both, morphological and molecular levels (Leuzinger *et al.*, 1998; Nagao *et al.*, 1998). Similarly, the ascidian *Otx* homologue is capable to rescue *otd* null mutants (Adachi *et al.*, 2001). Moreover, the fly *otd* is able to rescue most of the defects of *Otx1* and *Otx2* null mutant mice (Acampora *et al.*, 1998; Acampora *et al.*, 2001). These findings demonstrate that *Otx* homologues from diverse phyla have a high degree of functional equivalence in the homologous body regions, where they are normally expressed during development. This strongly supports the idea that the *Otx*/ *otd* function in the cephalic development program was present in a common ancestor of Deuterostomia and Ecdysozoa. Even though anterior neural expression of the *Otx* genes has also been reported in Lophotrochozoa (Bruce and Shankland, 1998; Umesono *et al.*, 1999), no rescue of *otd* phenotype by a

lophotrochozoan *Otx* homolog has been successfully carried out until now (Y.Adachi, unpublished). Therefore, the question, whether the observed functional equivalence between the ecdysozoan and the deuterostomian *Otx* homologues is conserved in lophotrochozoan ones, remains unanswered. Increasingly more elaborated brains, such as the vertebrate one, have been generated during evolution. This has been probably achieved by modifications in the spatial and/or temporal regulatory control of genes like *Otx*.

The functional equivalence of *Otx* and *otd* is probably due to their ability to activate conserved genetic pathways. Structural homology of Otx proteins is mainly limited to the homeodomain, which is likely to be responsible for their functional equivalence. To find out whether *Otx*/ *otd* regulate the same conserved downstream targets or different, but functionally equivalent, ones. Otx downstream targets have been sought for. No common target of *otd* and *Otx* has presently been identified, beyond a doubt (Boncinelli and Morgan, 2001). But cross-phylum overexpression experiments combined with gene expression analysis using microarrays have revealed some possible common downstream targets of the *D.melanogaster otd* and the human *hOtx2* (Montalta-He *et al.*, 2002). One third of the *otd*-regulated transcripts have been shown to also respond to human *hOtx2* overexpression in the fly. Those common targets probably illustrate the molecular basis of the functional equivalence of *otd* and human *hOtx2* in *Drosophila*.

1.2 The *Cdx*/ *cad* related genes

The *Cdx* gene, another homeodomain containing genes, has been shown to play a conserved role in the patterning of the A-P axis of bilaterian animals. It has been first identified in *D. melanogaster*, as the *caudal* gene (*cad*) (Mlodzik *et al.*, 1985). *Cdx* homologues share functions in the early patterning of the embryo and in the specification of the posterior body extremity in a wide range of animals. To date, *Cdx* homologues have been isolated from all of the three superclades: Deuterostomia, Ecdysozoa and Lophotrochozoa. In all species investigated, *Cdx* homologues have been shown to be expressed in the most posterior part of the body axis in the embryo (Pillemer *et al.*, 1998; Katsuyama *et al.*, 1999; Le Gouar *et al.*, 2003; Matsuo *et al.*, 2005).

1.2.1 The *Drosophila cad* gene

In *Drosophila*, the *cad* mRNA is maternally inherited and ubiquitously distributed in the unfertilized egg. Posterior to anterior gradients of maternal *cad* mRNA and Cad protein are rapidly forming during the syncytial blastoderm stage (Mlodzik and Gehring, 1987). Both, maternal and zygotic Cad gradients are required for specification of the posterior blastoderm and are involved in A-P axis patterning (Macdonald and Struhl, 1986). After gastrulation, the expression domain of *cad* flanks the one of the most posteriorly expressed *Hox* gene, *Abdominal-B* (*Abd-B*). The *cad* gene is essential for invagination of the hindgut primordium and its further specification and development. It regulates the expression of gap and pair-rule genes, which in turns regulate the *Hox* genes (Rivera-Pomar *et al.*, 1995). Cellular blastoderm embryos, which express ectopic *cad* at the anterior end, display defects in head development and segmentation. These defects are the results of alterations in the expression of segmentation genes, such as *fushi tarazu* and *engrailed*, and of the repression of head-determining genes (Mlodzik *et al.*, 1990). In adult flies, Cad is required for the development of the analia. This body region, which is derived from the most posterior body segment, is composed of the anal plates (external analia) and the hindgut (internal analia). The *cad* expression domain is restricted to the analia, abutting the one of *Abd-B*, the most posterior *Hox* gene. On one hand, *cad* down-regulates the expression of *Abd-B*, on the other hand, it induces the activity of downstream targets like *Distal-less*, *even-skipped* and *brachyenteron*, which are involved in the specification of the diverse regions of the analia. The two components of the analia, the hindgut and the anal plate, are specified by a combinatorial mode. While *cad* is required and sufficient for hindgut specification, it acts in combination with *Dll* and other genes to repress hindgut formation and consequently induce anal plate development (Moreno and Morata, 1999). In the absence of *cad* activity, the analia are transformed into male genitalia, which are normally developing from the segment immediately anterior to the *cad* expression domain. Furthermore, ectopic expression of *cad* can induce the development of analia in the head or the wing (Moreno and Morata, 1999). Hence, *cad* exhibits the behavior expected for a homeotic gene that specifies the most posterior part of the body. However, it is not physically linked to either the *BX-C* or the *ANT-C* loci and therefore is not considered

a *Hox* gene, per se. Nevertheless, in molecular phylogenetic analysis, the *cad* gene groups with the posterior *Hox* gene subfamily: the chordate PG9 to PG13 and the insect *Abd-B* (Brooke *et al.*, 1998). This suggests that *Cdx* related genes and the *Hox* genes might have arisen from a common ancestor. The discovery of the "ParaHox" cluster, firstly identified in *Amphioxus*, supports the concept of a "ProtoHox" cluster, from which the *Hox* and *ParaHox* genes have originated (Brooke *et al.*, 1998).

1.2.2 The vertebrate *Cdx* homologues

Three *Cdx* genes are known in vertebrates. The mouse *caudal* homologues, *Cdx1*, *Cdx3* and *Cdx4*, are expressed during gastrulation and neurulation in a graded manner along the A-P axis, with a preferential posterior up-regulation of their expression (Gamer and Wright, 1993; Northrop and Kimelman, 1994; Marom *et al.*, 1997). They are key players in a number of processes, including the early A-P patterning of the embryo, the specification of the posterior end, axial elongation and intestinal differentiation (Lohnes, 2003). In *Xenopus*, for example, *Xcad-2* partial loss-of-function mutant embryos develop enlarged heads and longer trunks. Conversely, *Xcad-2* overexpression generates embryos lacking head structures and exhibiting shorts axes (Epstein *et al.*, 1997). In contrast to *Drosophila*, where *cad* acts like a homeotic gene and specifies segment identity, vertebrate *Cdx* genes appear to act upstream of the *Hox* genes. Cdx members directly regulate vertebrate *Hox* expression in the mesoderm and the neurectoderm in a dosage-dependant manner via Cdx-binding sites, which are present in clusters throughout the *Hox* complexes (Charite *et al.*, 1998; van den Akker *et al.*, 2002). Specification and patterning of the posterior end of the embryo is dependent on a number of signaling molecules. Under appropriate conditions, the Wnt/ Wingless, the Fibroblast Growth Factor (FGF) members and the vitamin A metabolite retinoic acid (RA), are known to suppress anterior identity markers and induce expression of genes that are involved in posterior fates, such as certain *Hox* members (Altmann and Brivanlou, 2001). Some studies suggest that Cdx members may convey information from caudalizing agents, such as the Wnt, the FGF and RA, to the axial skeleton and the neurectoderm by regulating *Hox* gene expression (Lohnes, 2003).

1.2.3 *Cdx*-related genes and the axial elongation process

Another recently discovered conserved property of Cdx members is their involvement in the axial elongation process of bilaterian animals. This function is conserved in the crustacean *Artemia* and the beetle *Tribolium* (Copf *et al.*, 2004), in crickets (Shinmyo *et al.*, 2005), in the annelid *Platynereis* (de Rosa *et al.*, 2005) and in vertebrates (Chawengsaksophak *et al.*, 2004). In vertebrates, similarly to short-germ arthropods, somites are generated sequentially, from a posteriorly located presomitic zone, where *Cdx* members are expressed (Pourquie, 2003). Even though arthropod segments are ectodermal and vertebrate somites are mesodermal, molecular and morphogenetic similarities exist in the way that segmentation occurs in both, arthropods and vertebrates. By maintaining the self-renewing capacities of presomitic cells and by regulating *Hox* gene expression in the presomitic zone, Cdx members play important roles in axis elongation, somitogenesis and the specification of somite identity. Interactions between the Cdx members and signaling pathways, such as the FGF and the Wnt ones, are also important for the control of the axial extension by Cdx. Hence, in vertebrates, the Hox regulators Cdx are also integrated into a genetic network controlling A-P patterning, tissue generation and mesoderm segmentation (Fig. 2.2) (Deschamps and van Nes, 2005).

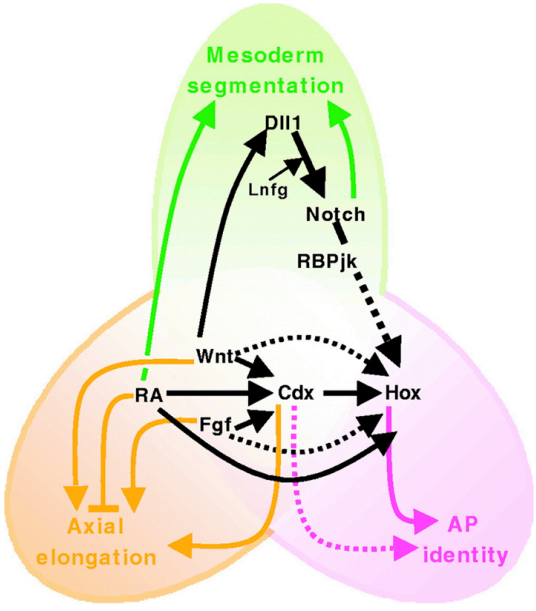


Fig. 2.2 Cdx members and the genetic network driving axial elongation, A-P patterning and mesoderm segmentation (Deschamps and van Nes, 2005)
Unbroken lines indicate established interactions; broken lines represent documented interactions, which have not yet been established at the molecular level.

1.3 *Otx*-related and *Cdx*-related genes, two ancestral key players of early A-P patterning in bilaterian animals?

1.3.1 Evidence from model organism studies

In a wide range of animals, Otx and Cdx, homeodomain-containing transcription factors, play a pivotal role in early patterning of the prospective head and the tail region, respectively. During early gastrula stages of *Xenopus* embryogenesis, *Otx2* and *Xcad3* are expressed in adjacent domains within the dorsal organizer, where they display mutual repression activities. Furthermore, interactions between *Otx2* and another *Xenopus caudal*-related gene, *Xcad2*, have also been described. They exhibit nested expression domains and are engaged in repressive regulatory interactions (Epstein *et al.*, 1997). In mid-gastrula stages, the embryo elongates and their expression domains separate, creating a gap, free of *Otx2* or *Xcad2* expression. In the median region, another homeobox gene, *Xgbx2a*, starts to be expressed. It is under negative regulation by both *Otx2* and *Xcad2*. This serves to establish the anterior and posterior expression boundaries of *Xgbx2a*. The sharp border between the *Otx2* and the *Xgbx2a* expression domains prefigures the location of the midbrain-hindbrain boundary (Simeone, 2000). These data argue that mutual repression of *Otx* and *Cdx* members is involved in setting up the early anterior and posterior domains within the *Xenopus* embryo (Isaacs *et al.*, 1999).

From such studies, it has been postulated that of a network of homeobox containing genes, such as *Otx* and *Cdx*, are acting in the early vertebrate embryo to subdivide the A-P axis into large domains. Subsequently, the pattern is refined by the activities of additional homeobox genes, like the *Hox* genes (Krumlauf, 1994). This situation evokes the body patterning and segmentation processes occurring in *Drosophila*, where large domains are first allocated and afterwards undergo refinement. Indeed, in the fly, the A-P axis is specified by unequal distribution of key maternal determinants within the egg cytoplasm. Maternal *bicoid* (*bcd*) transcripts are localized at the anterior pole of the egg. Its translation generates an anterior to posterior gradient of Bicoid protein. This morphogenetic gradient activates specific effectors genes required for segmentation, in a dosage-dependent manner (Driever and Nüsslein-Volhard, 1988): highest Bicoid concentrations activate head gap genes,

such as *otd*; lower dose activates the thoracic gap gene *hunchback* (*hb*) and the lowest levels activate the abdominal gap genes *krüppel* and *knirps*. Bcd and Hb cooperate in activating target genes involved in the specification of distinct head and thorax regions. Hb itself has morphogenetic properties and can create A-P polarity in the absence of Bcd (Hulskamp *et al.*, 1990). It specifies the anterior regions while preventing the formation of posterior ones. *Drosophila* embryos, lacking both maternal and zygotic *hb* mRNA, exhibit an almost bicaudal phenotype, in spite of the presence of *bcd* (Simpson-Brose *et al.*, 1994). Thus, in the absence of *hb, bcd* is not able to achieve correct polarization of the embryo. This finding emphasizes the crucial early patterning role of *hb*. Beyond its transcriptional regulator function, Bcd also represses the translation of the ubiquitously distributed maternal *cad* mRNA (Rivera-Pomar *et al.*, 1996). In this way, it creates a reciprocal posterior to anterior gradient of the Cad transcription factor. Later, pair-rule and segmentation genes subdivide the blastoderm into smaller regions via crossregulatory interactions, setting up the basic segments of the larvae. In *Drosophila*, the correct establishment of opposing morphogenetic gradients of the Hb and Cad proteins requires mutual repression between the anterior and posterior patterning systems. In *Drosophila*, the translation of the maternal *cad* mRNA is negatively regulated by Bicoid. Hence, the exclusion of the *Drosophila* Cad activity from anterior regions is accomplished in a different manner than the one observed in vertebrates. It is possible that *Drosophila* uses a derived developmental strategy: the negative transcriptional regulation of *Cdx* by Otx, described in *Xenopus*, might have evolved into a negative translational regulation of the maternal *cad* mRNA by the Bcd protein.

Only identified in cyclorrhaphan flies, *bcd* is a phylogenetically young gene that has evolved late (McGregor, 2005). It does not exist in other long-germ band insects, not even in the mosquito *Anopheles*, a close relative of flies. *bcd* and *zerknüllt* (*zen*), its sister gene, have probably evolved from duplication and mobility events of the insect *Hox3* gene at the stem of the cyclorrhaphan lineage (higher dipterans). The non-cyclorrhaphan flies (lower dipterans) possess a single *Hox3* gene, which displays a higher degree of similarity with *Drosophila zen* than with *bcd*. The orchestration of the A-P polarity by Bcd is therefore a developmental innovation, limited to cyclorrhaphan flies and might have evolved in concert with the fast development time of these flies (Stauber *et al.*, 2002; Lynch and Desplan, 2003) .

1.3.2 Evidence from *Nasonia* and *Tribolium* studies

Recently, the A-P patterning systems of other long-germ band insects, such as the hymenopteran *Nasonia*, and of short-germ band insects, such as the beetle *Tribolium*, have been investigated. In long-germ band insect embryogenesis, all segments are specified simultaneously within the blastoderm and are proportionally represented in the embryonic anlage. In contrast, in short germ-band insect embryogenesis, only anterior segments are patterned at the blastoderm stage and the remaining segments appear after gastrulation, sequentially generated from a growth zone (Davis and Patel, 2002). Hence, while the segmentation process in a long-germ band insect occurs by proportionate subdivision of the pre-gastrula anlage, segmentation in a short-germ band insect requires a phase of secondary growth, coupled with a sequential specification of the more posterior segments.

In *Tribolium*, in contrast to *Drosophila*, the *otd1* mRNA is maternally inherited and, as a result of translational repression from a posterior factor, Otd1 protein forms an anterior to posterior gradient. Removing *otd1* mRNA during oogenesis results in the loss of the entire head. This is a more severe phenotype than what is observed in *Drosophila otd* null mutants, where only the ocular and antennal segments are lost (Schroder, 2003). The *Tribolium otd1* mutant phenotype is reminiscent of a weak *Drosophila bcd* mutant phenotype (Frohnhöfer and Nüsslein-Volhard, 1986). As in *Drosophila*, *Tribolium hb* specifies posterior gnathal and thoracic segments. Furthermore, when both *otd1* and *hb* mRNA are removed during *Tribolium* oogenesis, the head, the thorax and the anterior abdomen of the embryo fail to develop (Schroder, 2003). This mutant phenotype is reminiscent of a strong *Drosophila bcd* mutant phenotype (Frohnhöfer and Nüsslein-Volhard, 1986). Hence, *otd1* and *hb*, not only specify the head and thorax, but also act synergistically during the segmentation process of the abdomen. Otd1 in combination with Hb probably plays the same transcriptional activator role like Bicoid in *Drosophila*. Bicoid and the Otx-related proteins both have a lysine (K) at the crucial position 50 of their homeodomain. The K50 of the homeodomain probably imparts the same DNA binding specificities to Otd and Bcd. Although Bicoid has been shown to repress *cad* translation in the anterior half of *Drosophila* eggs, Otd1 is not likely to mimic such a function in *Tribolium* eggs. The RNA-binding activity of Bicoid depends on the

presence of an arginine (R) at position 54 of the homeodomain. The homeodomain of Otx-related proteins exhibits an alanine rather than an arginine at this position (McGregor, 2005). Therefore, in contrast to Bicoid, Otx-related proteins cannot bind to RNA. Consequently, Otd1 probably cannot serve as a translational repressor of *cad* in *Tribolium.* A repressor of *Tribolium cad* is likely to exist and to serve as a further component of the anterior patterning system. Such a candidate repressor has not yet been identified in *Tribolium*. Different mechanisms of *cad* regulation have been described in other systems (Hunter and Kenyon, 1996; Isaacs *et al.*, 1999). Therefore, rather than a translational repression of *cad*, another kind of interaction may take place in *Tribolium*. However, a direct or indirect interaction with *otd1* cannot be excluded in the determination of the posterior expression domain of *cad*.

The *Nasonia* Otd1 also presents a Bicoid-like function in the anterior patterning of the embryo. As in *Tribolium*, *otd1* mRNA is inherited from the mother during oogenesis. Parental *otd1* inactivation by RNAi has produced headless embryos (Lynch *et al.*, 2006). The strong interaction between *otd1* and *hb* is crucial for the patterning of anterior body regions in *Nasonia*. A zygotic *hb* mutation deletes not only the thoracic and gnathal head segments, but also most of the pre-gnathal segments (Pultz *et al.*, 1999). A double knockdown of *otd1* and *hb* produces embryos that lack the entire anterior region, including several abdominal segments. This is a more severe phenotype than the combination of individual knockdowns of either *otd1* or *hb* (Lynch *et al.*, 2006). The *Nasonia* embryo relies more heavily on *hb* for the patterning of its anterior, than *Drosophila* or *Tribolium*. To date, *otd*/ *Otx* gene expression has been regarded as a marker for head development from annelids to mammals. Amazingly, in addition to its expression in an anterior to posterior gradient, *otd1* is also expressed in the posterior pole of *Nasonia* eggs and forms a posterior to anterior gradient, which has morphogenetic functions (Lynch *et al.*, 2006). Removing *otd1* parental mRNA in *Nasonia* deletes not only the head but also all segments posterior to abdominal segment 4. Therefore, anterior and posterior gradients of Otd1 proteins might have different functions: in combination with Hb, Otd1 is patterning the anterior end; lacking the synergy with the Hb cofactor, Otd1 activates distinct sets of target genes at the posterior end, possibly in combination with a different factor. Cooperation with the posterior expressed *Otd1* and *cad* has been suggested in *Nasonia* (Lynch *et al.*, 2006). These experiments raise the idea that *Drosophila* and

Nasonia have evolved divergent strategies to pattern their posterior segments. The maternal *Nasonia Otd1* could play some of the roles performed by the maternal *Drosophila cad* at the posterior end of the embryo (McGregor, 2006). However, the function of *cad* in *Nasonia* remains to be elucidated; its investigation will probably shed light on the evolution of the posterior patterning system in insects. Different interactions between *Otd1* and *Cad* have been suggested to occur in *Nasonia*: on one hand, a direct or indirect antagonistic interaction between *cad* and the anterior *otd1* gradient could take place. On the other hand, *cad* might synergistically interact with the posterior Otd1 gradient to pattern the posterior end of the embryo. Even though, this is highly speculative and requires support from additional functional data, finding such interactions would reinforce the idea that *Otx* and *Cdx* homologues are ancestral key players for early A-P body patterning. The mechanism, by which they interact, might have been redeployed during evolution to accommodate different patterning strategies depending on specific developmental constraints. It has been suggested that, by patterning all segments at once, *Drosophila* embryos are able to develop faster than short-germ band insects (Davis and Patel, 2002). This would have given a selective advantage by speeding up the development, especially in the transition from the short-germ band development mode to a long-germ band one in arthropods. The evolution of opposing Bcd and Cad gradients may have allowed a simultaneous anterior and posterior patterning to occur in a long-germ band embryogenesis.

Studies from *Nasonia* and *Tribolium* demonstrate that, in insects lacking *bicoid*, the anterior patterning is regulated by a combination of maternal *otd1* and *hb*. This suggests that a synergistic association between Otd1 and Hb may have performed the same role like Bcd in *Drosophila*. It is known that Hb cooperates with Bcd in *Drosophila*, with Otd1 in *Tribolium* and with Otd1 in *Nasonia*. However, the anterior organizing function of Hb has originated in the arthropod lineage and is therefore not likely to be conserved among bilaterian animals. The conserved functions of Hb within the Protostomia, and probably within the Bilateria, are actually the patterning of the epithelium and of the CNS (Pinnell *et al.*, 2006). Hence, *hb* was probably not a component of the ancestral A-P patterning system of the Deuterostomia and Protostomia ancestor.

Mutual repressions between early anterior and posterior domains, in part mediated by *Otx* and *Cdx*, seem to be part of a conserved mechanism of embryonic patterning. Studies from Deuterostomia and Ecdysozoa argue for evolutionary conserved mutual interactions between *Otx* and *Cdx* homologues in the patterning of A-P body axis. However, those inferences are based on data from only two of the bilaterian clades. To which extend their interactions are conserved and are part of an ancestral developmental mode remains to be clarified. Therefore, a comparative analysis of the A-P patterning system used by lophotrochozoan members, the third bilaterian super clade, should clarify this question.

1.4 *Otx*-related genes and the CNS development

Otx plays evolutionary conserved important roles in the development and specification of anterior structures, including the rostral regions of the developing brain of bilaterian animals. These findings also pose the question whether the nervous system is homologous across Bilateria.

1.4.1 Apparent dissimilarities in the CNS development of invertebrates and vertebrates

The central nervous system (CNS) of bilaterian animals is composed of two main parts, a nerve strand and an anterior cerebral center. In contrast to the nerve strand (ventral nerve cord in protostomes and spinal cord in vertebrates), the anterior cerebral nerve center, the "brain", is located at a similar position in all bilaterian animals, with the exception of the derived ones, such as bivalves or echinoderms. Based on traditional descriptive analyses of CNS development, Bilateria have been divided into two major groups: the Gastroneuralia and the Notoneuralia. The Gastroneuralia, such as arthropods and annelids, have a ventral nerve cord. The Notoneuralia, including all chordates, have a dorsal nerve cord. In gastroneuralian animals, the ganglionic masses separate from the ventral neurectoderm and form a ventral "rope-ladder" nervous system, containing connectives and commissures. In notoneuralian animals, the entire neurectoderm folds inwards to generate the dorsal neural tube (Fig. 2.3).

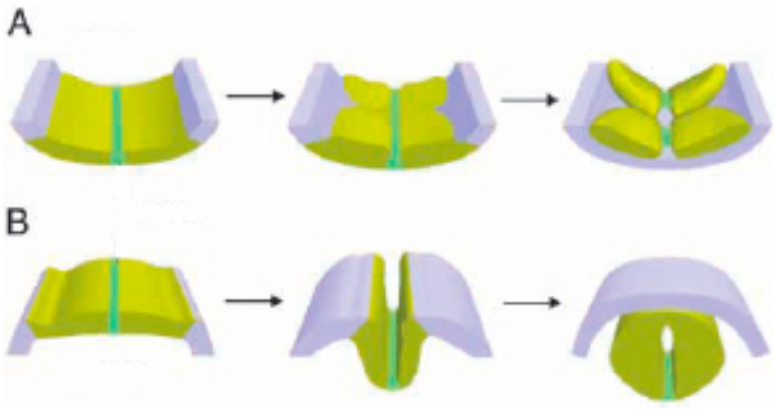


Fig. 2.3 Morphogenesis of the nerve strand in prototypic Gastroneuralia (A) and in prototypic Notoneuralia (B) (Arendt and Nubler-Jung, 1999)
The neurogenic ectoderm is in yellow-green (midline cells are in green). The epidermal ectoderm is in blue.

1.4.2 Homologous genes are involved in the CNS development of both invertebrates and vertebrates

Due to these dissimilarities in the morphogenesis and embryonic topography of the nervous system, it has been proposed that the nervous system has independently originated in the two groups, the Gastroneuralia and the Notoneuralia (Nielsen, 1995). However, molecular analyses of nervous system development in selected insect and vertebrate species have revealed astonishing similarities between the Gastroneuralia and the Notoneuralia in respect to the expression and function of regulatory genes, commonly involved in CNS embryogenesis. Indeed, many processes required for nervous system morphogenesis, such as patterning, proliferation, specification, and axonal outgrowth, have been found to be controlled in a comparable manner and by homologous genes in insects and vertebrates (Arendt and Nubler-Jung, 1999).

The early antagonistic activity between the secreted molecules of *decapentaplegic* (*dpp*) and *short gastrulation* (*sog*) and their homologues, respectively, subdivide the early ectoderm into epidermal and neurogenic ectoderm in both invertebrate and vertebrate species. The antineural function of *dpp* in invertebrates and *Bmp4* in vertebrates respectively, and the antagonizing neurogenic function of *sog*/ *Chordin* have been evolutionary conserved. But their site of

expression is inverted in respect to the dorso-ventral body axis. This finding has given strong support to the idea of a dorso-ventral axis inversion in the evolution of the deuterostomian branch. Originally postulated by Geoffroy Saint-Hilaire in 1822, this theory implies that the common ancestor of bilaterian animals, the Urbilateria, have already possessed a centralized nervous system, which has been inherited by both descendant lines, the Deuterostomia and the Protostomia. The CNS of both, vertebrates and invertebrates would therefore be homologous (De Robertis and Sasai, 1996).

Other levels of CNS specification are also conserved within the bilaterian animals: the position of cells becoming competent for neural fate in the early neurectoderm is controlled by "proneural" genes, encoding bHLH transcription factors, in both invertebrates and vertebrates. Among them are the *achaete-scute-complex*, the *neurogenin* and *atonal* homologues (Lee, 1997). By the process of lateral inhibition, accomplished by the "neurogenic" *Delta* and *Notch* homologues, the neural progenitor cells segregate from the non-neural ones (Lewis, 1996). Along the A-P axis, the neurectoderm of bilaterian animals is divided in different regions through the activity of a specific combination of conserved neural regionalization genes (Arendt and Nubler-Jung, 1996). In both insects and vertebrates, the developing nerve cord is subdivided in a *Hox*-expressing and a non-*Hox*-expressing region. While *Otx*/ *otd* and other "head gap" genes specify the anterior brain part, posterior brain parts, as well as the remaining nerve cord are regionalized by a similar expression pattern of *Hox* genes in the nervous system of the fly and the mouse (Hirth *et al.*, 1998).

Another level of the neurectoderm patterning is conserved: homologous genes, *vnd/ NKx2*, *ind*/ *Gsh-1* and *msh*/ *Msx,* are required for the dorso-ventral patterning of the developing CNS of Protostomia and Deuterostomia, respectively (Cornell and Ohlen, 2000). In *Drosophila*, proneural clusters and early delaminating neuroblasts are arranged in three longitudinal columns (medial, intermediate and lateral) on either side of the midline cells. In vertebrates, the proneural clusters and primary neurons likewise present this three-columns arrangement on either side of the neural plate. In accordance with the body axis inversion theory (De Robertis and Sasai, 1996), the expression of vertebrate homologues is inverted along the dorso-ventral axis

compared with that of their invertebrate homologues (Fig. 2.4). But, they are expressed in the same medial-lateral order: the homologous homeobox genes *vnd/ NKx2*, *ind*/ *Gsh-1* and *msh*/ *Msx* are expressed in a similar manner in the medial, intermediate and lateral neurogenic columns of both insects and vertebrates, respectively. They have been shown to be essential for the formation and specification of the neurogenic columns, in which they are expressed (Arendt and Nubler-Jung, 1999). It remains possible that these genes have been recruited independently during evolution for the D-V patterning of the CNS in both Protostomia and Deuterostomia. However, finding conserved expression of regionalization homologues in three longitudinal stripes, in correlation with neural progenitors arrangement in the corresponding columns, is a very specific pattern. It is unlikely that the same set of homologous genes has been recruited during evolution of the D-V patterning system of the CNS several times, independently. A common D-V patterning system of the CNS was rather already present in the common ancestor of both, Protostomia and Deuterostomia.

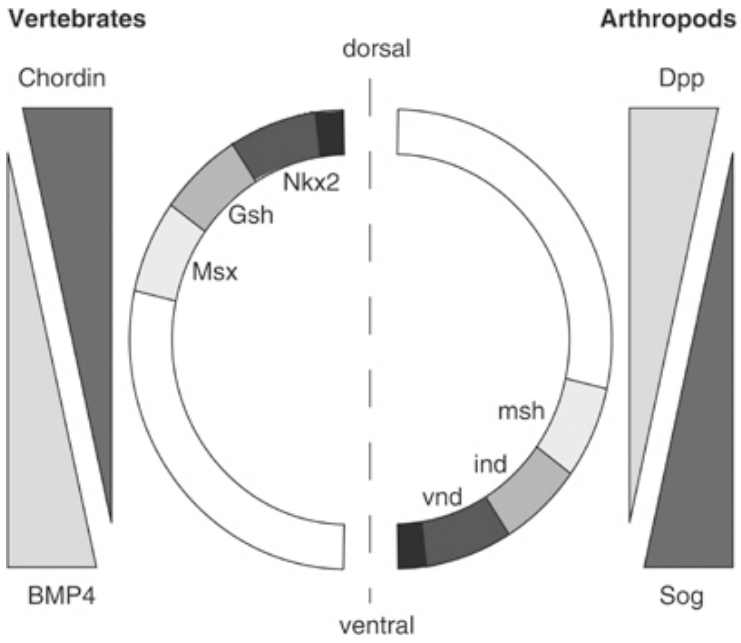


Fig. 2.4 Representation of the molecular patterning of the DV axis in vertebrates and arthropods (Lichtneckert and Reichert, 2005)
Only half of the body wall is represented in the scheme for both, arthropods and vertebrates, with dorsal side orientated to the top. *sog*/ *Chordin* homologues antagonize the *dpp*/ *Bmp4* homologues from the region of the embryo that will adopt a neurogenic potential: ventral side for arthropods *vs* dorsal one for vertebrates. Commonly shared set of homeobox genes patterned the specified neurogenic region into medial (*vnd*/*Nkx2*), intermediate (*ind*/*Gsh*) and lateral (*msh*/*Msx*) domains.

1.4.3 The tripartite brain organization hypothesis

As described above, the invertebrate and vertebrate brain is commonly subdivided into a rostral brain part, patterned by *otd*/ *Otx*, in a caudal brain part and the nerve cord, patterned by collinear *Hox* expression. In addition, a third brain-specific gene expression domain is conserved between the chordate and the arthropod brain. This has lead to the hypothesis that a tripartite organization of the brain has already been present in the common ancestor of chordates and arthropods (Reichert, 2005). The homeobox-containing *Gbx* genes function in brain development of both, *Drosophila* and vertebrates. They are expressed in a domain just posterior to the *otd*/ *Otx* domain. *otd*/ *Otx2* and *unpg*/ *Gbx2* homologues show mutual repression. This evolutionary conserved mechanism sets up the vertebrate midbrain-hindbrain boundary (MHB) and the *Drosophila* deutocerebral-tritocerebral boundary (DTB), which is homologous to the MHB. The MHB and the DTB are also defined as intermediate brain regions and are located between the anterior and the posterior brain regions, both in vertebrates and *Drosophila*, respectively. These intermediate brain regions are characterized by *Pax2/5/8* expression. They are positioned by the activities of otd/ Otx2 and unpg/ Gbx2, which are engaged in mutual repression. In vertebrates, it has been shown that either upregulation or downregulation of *Otx2* or *Gbx2* shifts the position of the MHB (Li and Joyner, 2001). The vertebrate MHB presents an organizer-like function, which has been first identified by transplantation experiments in chicken (Marin and Puelles, 1994; Rhinn and Brand, 2001). Although the expression of *Otx* and *Gbx* is essential to set-up and to maintain the MHB, *Otx* and *Gbx* are not required for the induction of MHB markers, such as *En*, *Wnt1* and *Pax2/5/8*. These markers are turned on later, independently of *Otx2* and *Gbx2* expression, and are responsible for the organizer properties of the MHB (Raible and Brand, 2004). Even though the fly homologues of the vertebrates MHB markers are expressed at the level of the fly DTB, the fly null mutants of the corresponding genes do not present brain-patterning defects (Hirth *et al.*, 2003). To date, there is no evidence for an organizer activity of the fly DTB as there is for the vertebrate MHB.

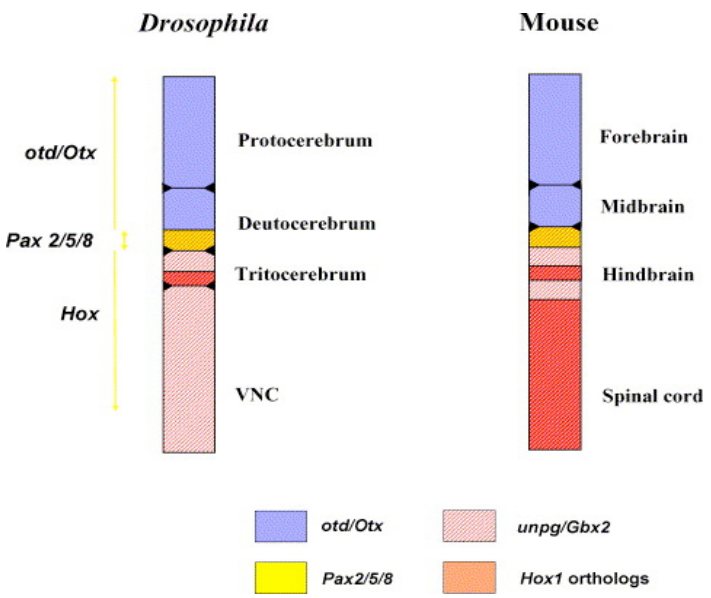


Fig. 2.5 The tripartite organization of the *Drosophila* and mouse brain based on the expression patterns of orthologous genes (Reichert, 2005)
Schematic representation of orthologues in the developing CNS of a *Drosophila* embryo at stage 13/14 and a mouse embryo at stage E10. In both, the *Pax2/5/8* expression domain in the embryonic brain is located between an anterior region expressing *otd/ Otx* and the posterior *Hox* expressing region and is positioned at the interface of the *otd/ Otx* expression domain and the posterior abutting *unpg/ Gbx2* one. The *otd/ Otx- unpg/ Gbx2* interface shows similar genetic features in *Drosophila* and mouse.

Also in Amphioxus, the anterior *Otx* and the more posterior *Gbx* expression domains abut at the boundary between the cerebral vesicle and the hindbrain, where a mutual antagonistic activity between *Otx* and *Gbx* has been suggested (Castro *et al.*, 2006). A possible homology for the Amphioxus boundary between the cerebral vesicle and the hindbrain with the vertebrate MHB is suggested from *Otx* and *Gbx* expression domains. However, the genes homologous to the MHB markers, which provide the organizer properties, are not co-expressed in this Amphioxus region. Therefore, it is unlikely that this border region, between the cerebral vesicle and the hindbrain, possesses organizer activity in Amphioxus. Thus, the apposition of the *otd/ Otx2* and *unpg/ Gbx2* expression domains, which specifies an evolutionary conserved intermediate brain region, may pre-date the acquisition of an organizer-like function by the MHB, since an organizer activity of this intermediate brain region has been described beyond a doubt only in vertebrates. In order to determine where in the chordate lineage an organizer function of the MHB has appeared, the tunicates (also named urochordates) have been investigated. However, it is difficult to resolve the

question of whether the tunicates have secondarily lost a MHB organizer activity or have never possessed one (Castro *et al.*, 2006). It is indeed complex to reconstruct the ancestral tunicate as many genes, including *Gbx*, have been lost by tunicates, during evolution. In addition, very different expression patterns of homologues genes, such as *en* and *Pax2/5/8*, are observed among tunicates. Finally, the phylogenetic position of tunicates is unclear: while they have traditionally been placed at the base of chordates, recent phylogenetic analysis using a large number of genes suggest that tunicates are a sister group to vertebrates and that Amphioxus is actually basal within chordates (Delsuc *et al.*, 2006). Thus, the question of where a MHB organizer activity has evolved in the chordate lineage is still under debate.

The nervous systems of bilaterian animals share many essential developmental features. A general homology of brains is broadly considered a reasonable hypothesis, at least in the "molecular field". Both, the vertebrate and invertebrate nervous systems are probably derived from a common ancestor that already possessed a rather sophisticated one. However the level of complexity already acquired in the urbilaterian CNS remains unclear. There is little doubt that the *Otx*- *Pax*- *Hox*-subdivision of the vertebrate brain is homologous to the similar subdivision reported in flies. The tripartite regionalization of the neural tube is very ancient. However, additional data from other phyla are required to test the idea that a tripartite ground plan of the CNS was already present in the common ancestor of the bilaterian animals, the Urbilateria. It is therefore necessary to get more information from the expression patterns of the *Otx*- *Pax*- *Hox*- homologues in a wider range of protostomes, including the Lophotrochozoa.

2. Aim of the first project of this PhD thesis

One of the most important breakthroughs in the field of developmental biology has been the discovery of the homeobox and of its widespread phylogenetic conservation. These findings make the first link between molecular data and bilaterian body plan architecture. The so-called *Hox* genes are known to play a pivotal role in the antero-posterior specification of bilaterian animals. Also other homeobox-containing genes play evolutionary conserved roles in the specification of the most rostral and the most caudal regions of bilaterian embryos: For example, *Otx* genes are required for the development of the most rostral part of bilaterian embryos. *Cdx* genes have been proposed to be required for the development of the most caudal part. Studies from Deuterostomia and Ecdysozoa suggested evolutionary conserved interactions between *Otx* and *Cdx* genes. These interactions are thought to discriminate the anterior from the posterior domains very early in the embryonic development of Bilateria.

During this PhD work, we wanted to investigate the specification of the nemertean A-P body axis during development and regeneration to find out whether the rostral/ caudal specification mechanism is not only conserved in Deuterostomia and Ecdysozoa but also in Lophotrocohozoa, the third clade of the Bilateria. The identification of several *Ls-Hox* genes is already in favor of the probable conservation of the A-P patterning mechanism in *Lineus* (Kmita-Cunisse *et al.*, 1998). However, it is not instructive for the very early discrimination between anterior and posterior body domains and for the specification of the body extremities. Therefore we have decided to search for *Otx* and *Cdx* homologues. We thought to test whether the presumed roles of *Otx* and *Cdx* in the A-P body axis specification and in development of the CNS (especially in the case of *Otx*) are conserved in a lophotrochozoan animal. To test this hypothesis, we asked the following questions:

- Are *Otx* and *Cdx* homologues involved in the specification of the most anterior part and the most posterior part, respectively, of a *Lineus* larva?

- If so, can we use these genes as markers to distinguish anterior regenerating parts from posterior ones in very early regenerative stages of *L.sanguineus*? Since there is no morphological feature that would allow us to recognize their anterior end from their posterior one.
- As suggested from studies in Deuterostomia and Ecdysozoa, *Otx* is playing an evolutionarily conserved role in the development of the *Lineus* CNS? If so, is it specifically expressed during its regeneration, too?

Because the presence of some original nerve cord parts is the only requirement for a complete regeneration in *L.sanguineus*, it has been suggested that, in response to body injury, the *Lineus* CNS could either emit signals responsible for regeneration, or integrate such signals or both. *Otx*, if expressed during *Lineus* CNS regeneration, could be part of such signaling network responsible for *L.sanguineus* regeneration. A difference in the ability of emitting and/or integrating signals of the nemertean CNS could be responsible, at least in part, for the different capacities of regeneration that are found among nemerteans. We wondered whether these different regeneration capacities are reflected in differences of activation of a gene such as *Otx*. Therefore, we asked:

- Are there any differences in the *Otx* expression patterns in amputated *Lineus* species that differ in their regenerative capacities, such as *L.sanguineus* and *L.lacteus*?

To approach such questions, we cloned the *Otx* and *Cdx* homologues from *L.sanguineus*. Then, we investigated their expression patterns during development, adult life and during different types of regeneration. We used ISH and real-time PCR approaches in order to monitor more precisely where and when those genes are up- or down-regulated in these processes (See Results).

3. Results

3.1 Identification of *Otx* and *Cdx* homologues from *Lineus sanguineus*

3.1.1 Cloning of *Ls-Otx*:

In order to clone *Ls-Otx*, we searched, in databases like Genbank, for the availability of *Otx* sequences from various animals. Aligning amino acid sequences of Otx homologues led to the identification of highly conserved regions. Several degenerated primers were designed based on these conserved regions. Different combinations of primers and conditions were tested to amplify an *Otx* homologue from *L.sanguineus*. The successful combination was a first PCR with primers Otx-Forward1 (corresponding to the sequence PRKQRRER) and Otx-Reverse1 (sequence WFKNRRA) followed by a nested PCR with primers Otx-Forward2 (sequence RRERTTFT, overlapping the sequence of the Otx-Forward1 primer) and Otx-Reverse2 (sequence MREEVALKIN). A 77 base pairs (bp) fragment of the homeobox of a putative *Otx* homologue was obtained and was subsequently extended by 3' and 5' RACE PCR procedures, using specific primers. Sequencing of the RACE PCR fragments allowed us to identify this gene as a *Ls-Otx* gene, based on its close similarity with other *Otx* genes. In comparison with *Otx* homologue sequences, the open reading frame (ORF) of *Ls-Otx* is complete. We have also obtained part of the 5' untranslated terminal region (UTR).

The *Ls-Otx* sequence contains an ORF of 936 bp, which encodes a putative protein of 312 amino acids (aa) (Fig. 2.6). The putative Ls-Otx protein contains a homeodomain sequence, which exhibits a glutamate at position 4 (E4) and a lysine at position 50 (K50). From phylogenetic analyses, it appears that the *Lineus* homeodomain groups with members of the Otx family, and not with other paired-class homeodomains (see appendix 1.). The K50, characteristic for Otx-type homeodomain, is crucial for its DNA-binding capacity and confers specificity for the TAATCC/T sequence (Hanes and Brent, 1989). In addition to the homeodomain, Ls-Otx contains another highly conserved region, a so-called WSP motif, but lacks a clear C-tail motif of 6 conserved amino acids, present in many Otx sequences.

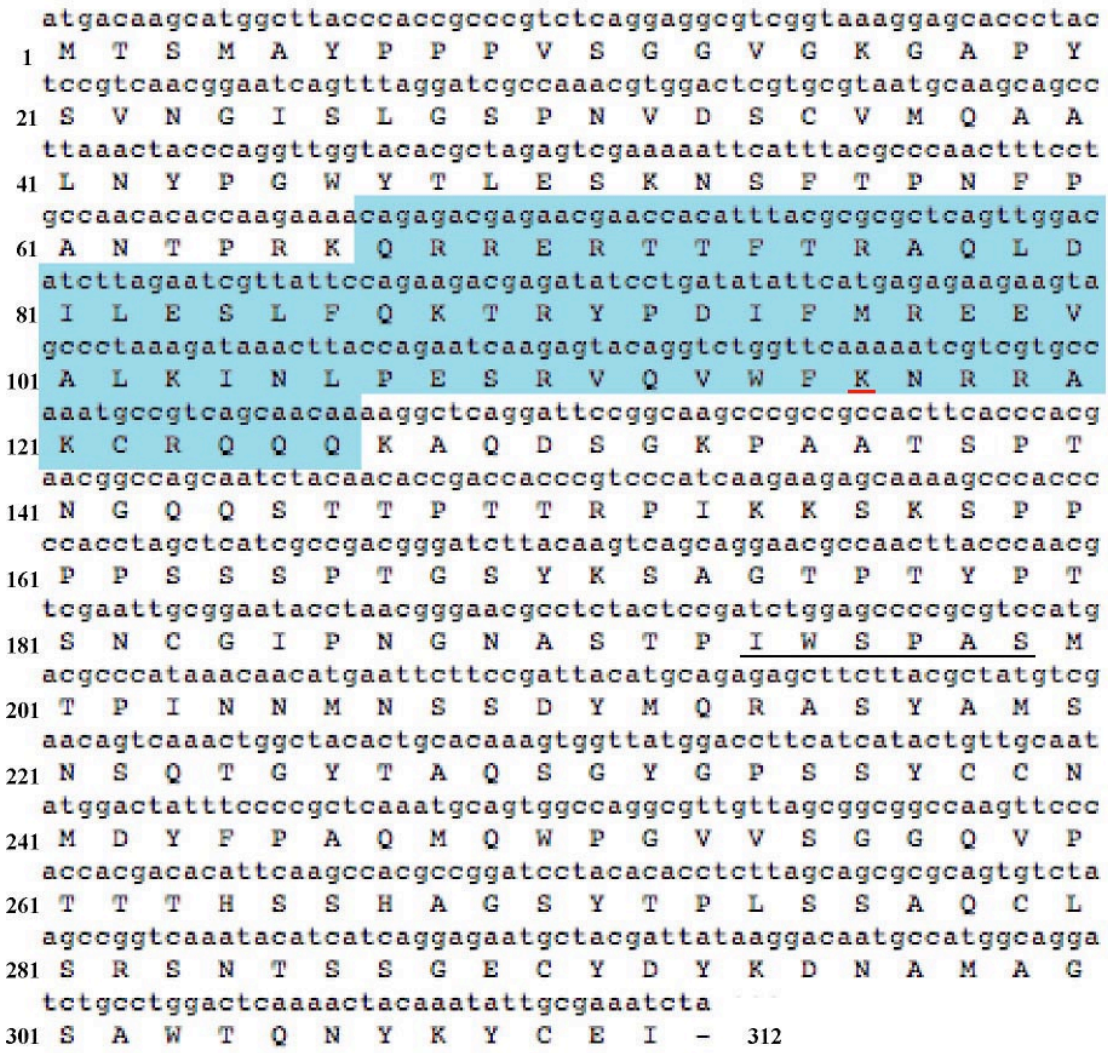


Fig. 2.6 The *Ls-Otx* ORF and its deduced aminoacid sequence
The homeobox sequence and its deduced aminoacid sequence are highlighted in blue; note the presence of the K50, underlined in red, within the homeodomain. Conserved aa residues of the WSP motif are underlined.

Using the same strategy, we cloned a small fragment of 159 bp of an *Otx*-homeobox-like sequence from a related species, *Lineus viridis*. As expected, the conservation of the homeodomain sequence is very high between *L.sanguineus* and *L.viridis* (Fig 2.7). Remarkably, sequence analysis revealed more than 94% identity at the nucleotide level.

```
94.3% identity in 159 residues overlap

Ls-Otx       190 CCAAGAAAACAGAGACGAGAACGAACCACATTTACGCGCGCTCAGTTGGACATCTTAGAA
Lv-Otx         1 CCCCGGAAACAGCGACGGGAACGTACCACATTTACGCGCGCTCAGTTGGACATCTTAGAA
                 **  * ****** **** ***** ************************************

Ls-Otx       250 TCGTTATTCCAGAAGACGAGATATCCTGATATATTCATGAGAGAAGAAGTAGCCCTAAAG
Lv-Otx        61 TCGTTATTCCGGAAGACGAGATATCCTGATATATTCATGAGAGAAGAAGTAGCCCTAAAG
                 ********** ************************************************

Ls-Otx       310 ATAAACTTACCAGAATCAAGAGTACAGGTCTGGTTCAAA
Lv-Otx       121 ATAAACTTACCAGAATCAAGAGTACAGGTATGGTTTAAA
                 *************************** ***** ***

98.1% identity in 53 residues overlap

Ls-Otx        64 PRKQRRERTTFTRAQLDILESLFQKTRYPDIFMREEVALKINLPESRVQVWFK
Lv-Otx         1 PRKQRRERTTFTRAQLDILESLFRKTRYPDIFMREEVALKINLPESRVQVWFK
                 ********************** ******************************
```

Fig. 2.7 Alignment of the *Lv-Otx* PCR fragment obtained by degenerated PCR and the *Ls-Otx* homeobox corresponding sequence
The sequence identity at the nucleotide level is 94,3% and 98,1% at the amino acid level. (Sequences covered by the degenerated primers are not included in the alignment).

3.1.2 Cloning of *Ls-Cdx*:

A partial 378 bp fragment from a *Cdx* homolog was already available from *L.sanguineus* (Kmita M., PhD thesis, 1995), when I started my thesis work. We designed specific primers in order to clone the full-length sequence by RACE PCR. We obtained a fragment of *Ls-Cdx*, which contains an ORF of 807 bp encoding a putative protein of 269 aa (Fig. 2.8). We also cloned some 3' and 5' UTR from *Ls-Cdx*. The putative Ls-Cdx protein contains a Cdx-type homeodomain. It also exhibits a hexapeptide motif specific for Cdx-type proteins (Marom *et al.*, 1997). From phylogenetic analyses, Ls-Cdx clearly groups with members of the Cdx family (see appendix 1.).

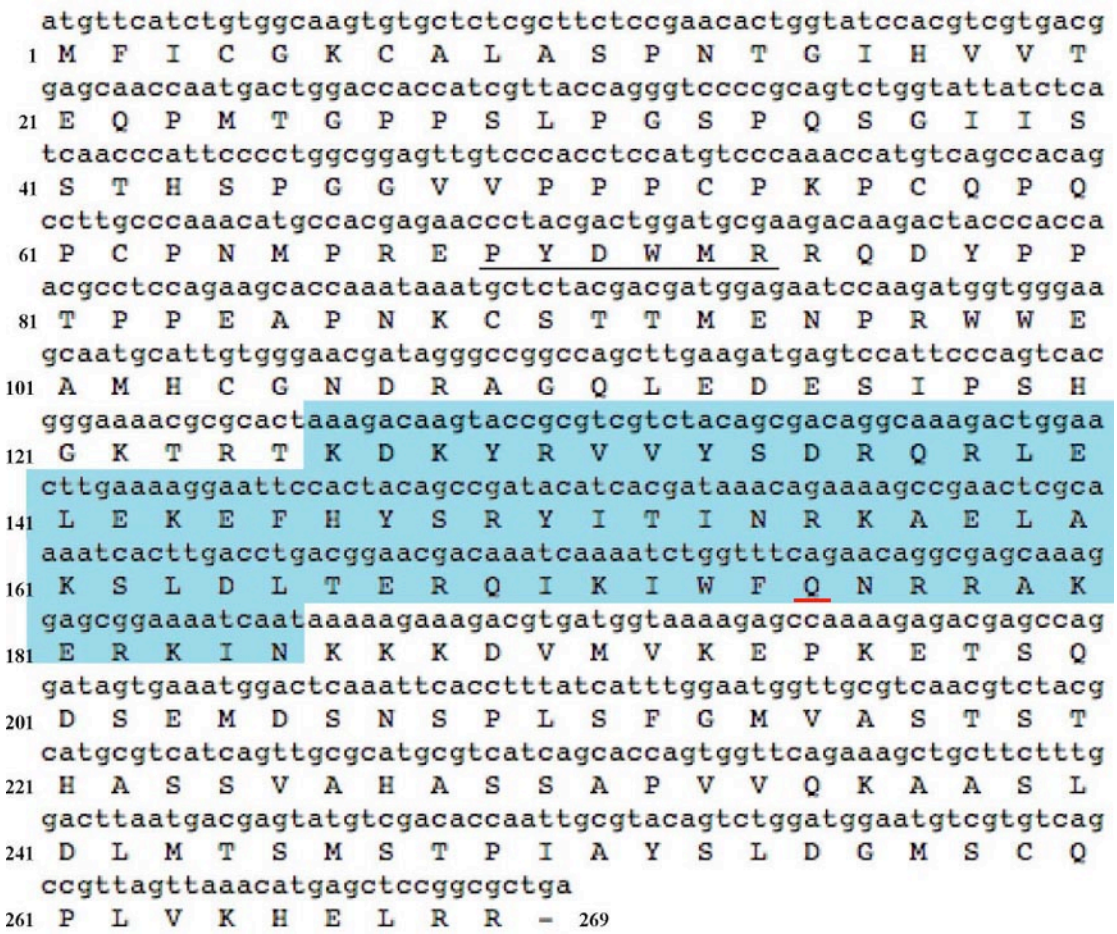


Fig. 2.8 The *Ls-Cdx* ORF and its deduced aminoacid sequence
The homeobox sequence and its deduced aminoacid sequence are highlighted in blue. Note the presence of the Q50, underlined in red, within the homeodomain. Conserved aa residues of the hexapeptide motif are underlined.

3.2 Expression patterns of *Lineus Otx* and *Cdx* homologues during development

3.2.1 The use of *L.viridis* to investigate *Lineus* development

As *L.sanguineus* does not sexually reproduce in captivity, *Otx* expression during embryonic development could not be examined in *L.sanguineus*. Therefore, to examine *Otx* and *Cdx* expression patterns during development of *Lineus*, we used a closely related species, from which we can easily collect larvae: *Lineus viridis.* *L.sanguineus* and *L.viridis* are close related species. Therefore, we made the assumption that an *Otx* homologue from *L.viridis*, if present in its genome, would probably have a homeobox sequence close to the one of *L.sanguineus*.

Actually, we cloned, in a similar approach than the one used for cloning *Ls-Otx*, a small *Otx*-like homeobox fragment from *L.viridis*. The sequencing of this *Lv-Otx* fragment revealed an extremely high degree of identity, at the nucleotide level, between *Ls-Otx* and *Lv-Otx* homeoboxes (Fig. 2.7). This suggests that these two homeobox sequences are sufficiently conserved to allow cross-hybridization of a *L. sanguineus* probe to the *L.viridis* transcript, under highly stringent conditions in *in situ* hybridization (ISH) experiments. By analogy, we made the same assumption for *Ls-Cdx* and *Lv-Cdx* homeobox sequences. Therefore, we examined *Lv-Otx* and *Lv-Cdx* expression patterns in developing *L.viridis* worms by using Dig-labeled anti-sense RNA probes comprising only the homeobox sequence of *Ls-Otx* and *Ls-Cdx*, respectively.

3.2.2 *Otx* expression pattern during *Lineus* development

In situ hybridization on young, 12 days old *L.viridis* larvae (Fig. 2.9, A2) revealed a strong *Otx* expression in the entire developing CNS. Both ventral cerebral ganglia, connected by the ventral commissure, and dorsal cerebral ganglia, connected by the dorsal commissure, are strongly stained. Moreover, *Otx* seems to be expressed in the pair of developing lateral nerve cords, at the level of their emergence from the ventral cerebral ganglia of the brain. The developing cerebral organs, which are sense organs that contain neurosecretory cells, present also a high level of *Otx* expression. These organs consist of a pair of invaginated epidermal canals, whose spherical inner end is embedded in a mass of glandular and nervous material, fused with the dorsal cerebral ganglia.

A weak *Otx* expression can also be observed in the developing gut. Nemerteans are known to exhibit strong endogenous phosphatase activity in their gut. This is often a challenging technical problem for scientists who want to perform ISH on these worms. However, we should notice that there is probably no or not that much endogenous phosphatase activity in this young developmental stage of *L.viridis*, as larvae hybridized with sense probe do not show such staining.

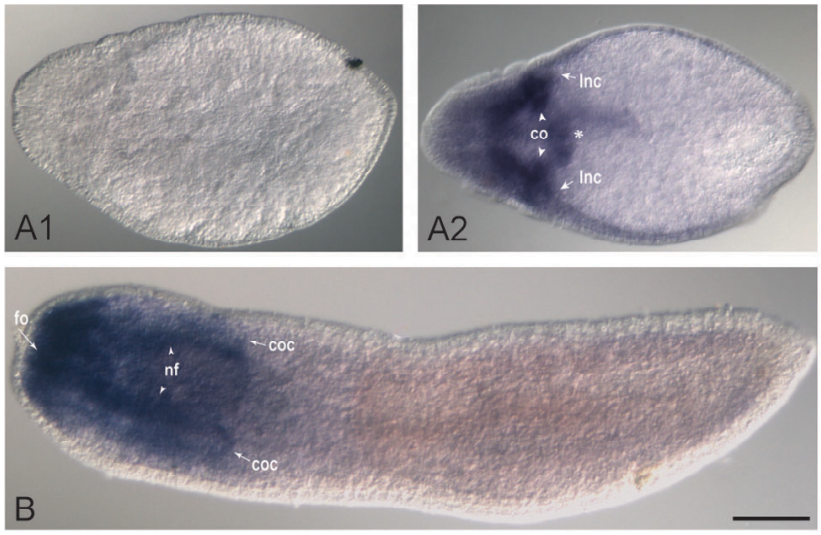


Fig. 2.9 Expression analysis of *Ls-Otx* in developing *L.viridis* worms
Whole-mount ISH were done using *Ls-Otx* sense (A1) and anti-sense (A2 and B) RNA probes. Anterior is oriented to the left, dorsal view. Age of the larvae is according to the number of days post cocoon discovery in our *L.viridis* laboratory stock.
co: cerebral organs; lnc: lateral nerve cords; fo: frontal organ; nf: nerve fibers; coc: cerebral organs canals. Asterisk: expression in the developing gut. Scale bar: 300 µm.
A: 10 days old larvae. B: 18 days old larvae.

Later in development (Fig. 2.9, B1), *Otx* is expressed in another developing sense organ, the frontal organ, and in the nerve fibers, connecting this organ to the cerebral ganglia. The expression of *Otx* in the brain and cerebral organs is weaker than it was in earlier developmental stages. But the cerebral canal, joining the cerebral organs to the external environment, still clearly expressed *Otx*. In addition, there is a global diffuse staining of all the anterior part of the *L.viridis* larvae. Finally, in contrast to the youngest developing stages, the lateral nerve cords and the gut do not expressed *Ls-Otx*.

3.2.3 *Cdx* expression pattern during *Lineus* development

As revealed in figure 2.10, *Ls-Cdx* is strongly expressed at the posterior end of the developing *L.viridis* larvae. Being rather internal, the staining seems to be at the level of the endoderm, in the developing intestine. In contrast to *Ls-Otx* expression, *Ls-Cdx* seems to conserve the same expression pattern during development, as it is restricted to the extreme posterior end of the developing intestine.

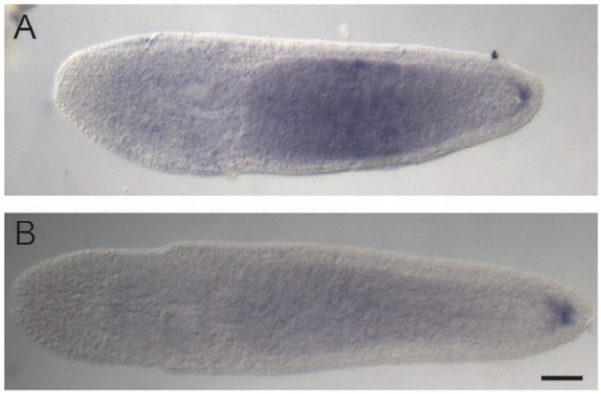


Fig. 2.10 Expression analysis of *Ls-Cdx* in developing *L.viridis* worms
Whole-mount ISH were done using *Ls-Cdx* anti-sense RNA probe.
Anterior is oriented to the left, dorsal view. Age of the larvae is according to the number of days post cocoon discovery in our *L.viridis* laboratory stock.
Scale bar: 220 µm for A and 340 µm for B.
A: 15 days old larvae. B: 22 days old larvae.

3.3 Expression patterns of *Ls-Otx* and *Ls-Cdx* in adults and during regeneration

To investigate *Ls-Otx* and *Ls-Cdx* expression in adult *L.sanguineus* and during regeneration, we synthesized, for each gene, a sense RNA probe and an anti-sense RNA probe, excluding the homeobox sequence. By doing so, we wanted to avoid possible cross-hybridization of the probes with other homeobox sequences.

3.3.1 *Ls-Otx* expression in adults *L.sanguineus*

In adult *L.sanguineus*, *Ls-Otx* expression remains restricted to the anterior part of the worm (Fig. 2.11). It is still expressed in the CNS of the adult, where it is detected at the periphery of cerebral ganglia, on their external side. In contrast to larvae, *Ls-Otx* is not detected in the lateral nerve cords in adult worms. Staining is also associated with sense organs: *Ls-Otx* is expressed all along the cephalic slits and specifically at the level of the pores of the cerebral organs. The cephalic slits are deep lateral grooves, lined with a modified epithelium. Sensory ciliated cells, lying over a zone of ganglia cells, are present in the cephalic slit modified epithelium. The ganglion cell area and the overlaying cephalic slit epithelium, which are both assumed to possess chemotactic functions, are expressing *Ls-Otx*.

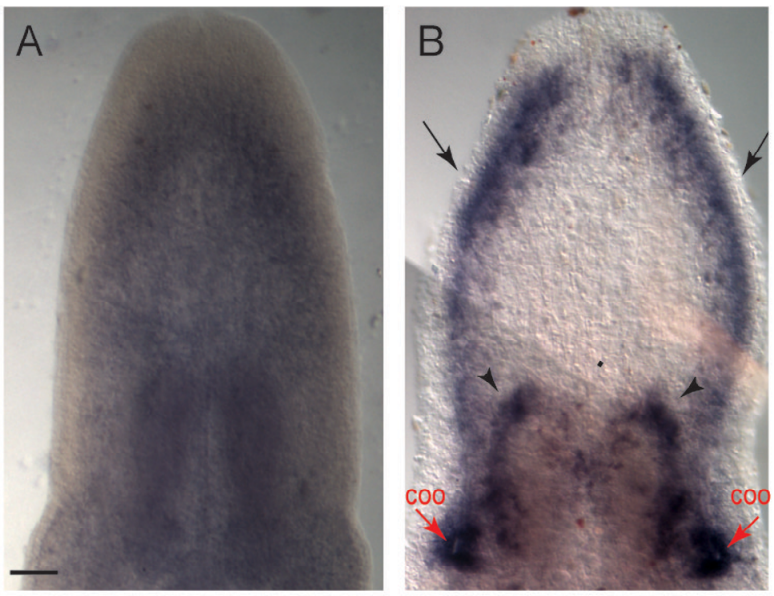


Fig. 2.11 *Ls-Otx* expression in an adult *L.sanguineus* head
Dorsal view, anterior is oriented to the top. Scale bar: 450 µm. coo: cerebral organs openings
A: *Ls-Otx* sense RNA probe
B: *Ls-Otx* anti-sense RNA probe. Expression in the external part of the cerebral ganglia is shown by arrowheads. Red arrows indicate expression at the level of the opening of the cerebral organ canal to the external environment is shown. Note that we can distinguish the opening of the canal (= the cerebral organ pore) as an unstained spot in the middle of a circle of cells strongly expressing *Ls-Otx*. Black arrows highlight the regular expression all along the cephalic slits running on both sides of the worm from the top of the head to the cerebral organs openings.

Like in the developing *L.viridis* worms, *Otx* expression is observed in the cerebral organs of adult *L.sanguineus*. The highest level of expression seems to be at the level of the pores, where the cerebral organs can communicate with the external environment and release viscous and refringent secretions. The precise function of cerebral organs is not clearly understood: they have been associated with a plethora of sensory functions such as audition, respiration, chemotaxis for the detection of food and/ or the analysis of water and finally associated with some endocrine function (Gibson, 1972).

3.3.2 *Ls-Cdx* expression in adults *L.sanguineus*

Even though we tried different *in situ* hybridization conditions (by changing hybridization temperature, by increasing hybridization time and even by making new RNA probes spanning different regions of the *Ls-Cdx* gene), we didn't detect any significant staining in adult *L.sanguineus*, which have been subjected to *Ls-Cdx* ISH.

We just obtained sometimes weak staining in the intestine of adult worms. This probably corresponds to background due to endogenous alkaline phosphatase activity.

3.3.3. *Ls-Otx* expression during regeneration of *L.sanguineus* antecerebral end

The adult brain and the anterior sensory organs of *L.sanguineus* strongly express *Ls-Otx* (Fig. 2.11). Therefore, we took advantages of the regeneration capacities of *L.sanguineus* to test whether *Ls-Otx* participate in regeneration of the brain and of the antecerebral sense organs. Using a razor blade, we made a cut at the boundary between region 3 and region 4 of *L.sanguineus* (Fig. 2.12, A). Such amputated worms were kept in good conditions (in the dark with several sea water changes) to allow the regenerative process to take place. 12 days afterwards, when a new brain is regenerating from the blastema, worms were fixed and subjected to whole mount *in situ* hybridization.

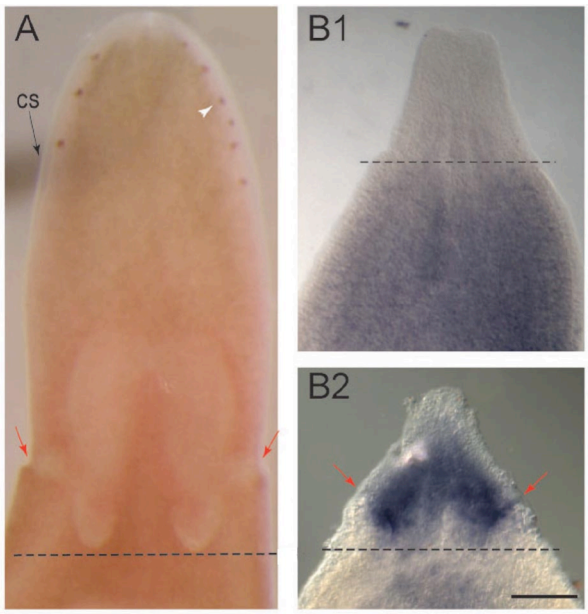


Fig. 2.12 *Ls-Otx* expression during anterior regeneration of *L.sanguineus* brain
Dorsal view, anterior is oriented to the top. Scale bar: 650 µm. cs: cephalic slit.
A: antecerebral head of an adult *L.sanguineus*. Notice the presence of numerous eyes (arrow head). The cephalic slits are visible on the left side on the head (black arrow). Red arrows point at the openings of the cerebral organ canal to the external environment.
The dashed lines indicate the amputation plane.
B1 and B2: ISH has been performed with, *Ls-Otx* sense (B1) and *Ls-Otx* anti-sense (B2) RNA probes on regenerating worms, fixed 12 days after the amputation. Dashed lines indicate the level of the amputation; tissues in front of the lines are newly regenerated ones.

Ls-Otx is specifically expressed in the entire regenerating brain: in both dorsal and ventral ganglia and their commissures (Fig. 2.12, B2). Diffuse staining is also observed in the cephalic nerves originating from the anterior surfaces of the dorsal ganglia. Regenerated cerebral organs continue to express *Ls-Otx*. Interestingly, these sense organs are the first to regenerate after cephalic amputation. *Ls-Otx* expression starts as soon as they start to be regenerated from the blastema.

The antecerebral end of the worm, also called body region 1, contains the cephalic glands and the eyes (Fig. 1.6). Interestingly, this is the only worm module that completely lacks CNS nerve cell bodies. If the antecerebral end of the worm is surgically removed, the remaining fragment is capable of regeneration and of restoring the normal body pattern. In fact, *L.sanguineus* has the great ability to give rise to an entire new worm from almost any piece of its body. However, an isolated body region 1 cannot undergo regeneration as it lacks CNS nerve cell bodies and therefore, degenerates (see Fig.2.25). Because of the presumed important role of CNS during *Lineus* regeneration (Coe, 1932), we decided to look for *Ls-Otx* expression during regeneration of only region 1. To do so, we amputated the worm just in front of the cerebral ganglia. In this situation, regeneration occurs rapidly, within few days. We didn't detect any *Ls-Otx* expression in the blastema of such regenerating worms (not shown). The *Ls-Otx* expression in the brain and the cerebral organs was detectable, as expected from ISH results performed in non-regenerating adults.

3.3.4 *Ls-Cdx* expression during posterior regeneration of the *L.sanguineus* gut

As *Ls-Cdx* is expressed in the developing intestine (Fig. 2.10), we decided to test whether it is also expressed during intestinal regeneration. Actually, *Ls-Cdx* seems to be expressed during the regeneration of the entire gut: it is upregulated during posterior regeneration of the gut from either the esophagus region (Fig. 2.13, A) or the intestinal region (Fig. 2.13, B). It exhibits a precise endodermal expression pattern at the level of the most posterior part of the regenerating gut.

Interestingly, we can observe that *Ls-Cdx* is only expressed at one extremity of the isolated intestinal fragment that undergoes both, anterior and posterior regeneration (Fig. 2.13, B).

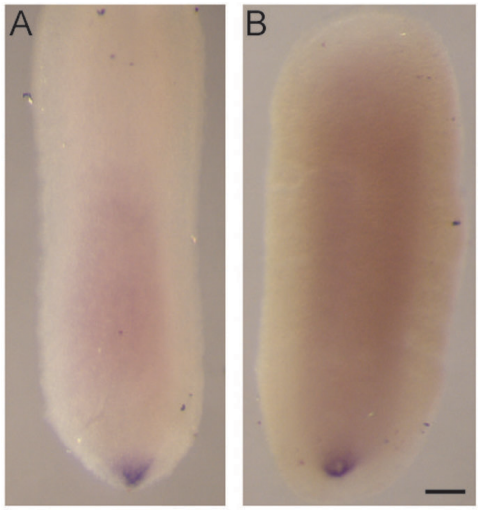


Fig. 2.13 *Ls-Cdx* expression during posterior regeneration of *L.sanguineus* gut
ISH were done using *Ls-Cdx* anti-sense RNA proble. Anterior is oriented to the top from sample A. Samples were fixed with paraformaldehyde 7 days after amputation or section. Scale bar: 450 μm in A and 300 μm in B.
A: adult worm, amputated from body regions 7 to 10, undergoing posterior regeneration.
B: intestinal fragment (region 8) isolated from an adult worm.
NB: worm from A was decapitated immediately before paraformaldehyde fixation. Therefore, the head region is missing on the picture.

3.3 Real time PCR approach:

3.3.1 General considerations of the real time PCR procedure

Real-time PCR has been developed because of the need to quantify differences in mRNA expression levels. Our assay is based on the detection and quantification of the fluorescence emitted during the reaction as an indicator of amplicon production at each PCR cycle ("in real time"). We used SYBR green, a dye that fluoresces very brightly when bound to double strand DNA. This technique measures total DNA synthesis; therefore we need to make sure that we are measuring a signal due to our real amplification target and not due to an artefact, like for example a primer-dimer. Extensive optimisation of the real-time PCR conditions is therefore required.

At the end of each run, the real-time PCR instrument determines the melting point of the PCR product. This is an important mean of quality control, as we can verify that all samples have a similar melting point. This allowed us to make sure that

real-time PCR artefacts, such as primer-dimer formation, primer mispriming and genomic DNA contamination, are absent from our experiments. Furthermore, by running a classic agarose gel, it can be verified whether the real-time PCR product has the expected size. All these quality control point permits us to obtain reliable results. In order to normalize the fold change in the expression of our target gene relative to "reference gene", we need an internal standard. We therefore cloned a fragment of *Ls-EF1α* (Elongation Factor 1 alpha) to be used as a reference gene. We made the assumption that *EF1α* expression level doesn't change significantly when *L.sanguineus* are undergoing regeneration. So, *EF1α* signals were used to normalize our target genes values: fold change values in target gene (*e.g Ls-Otx* or *Ls-Cdx*) were divided by *Ls-EF1α* fold change values.

In all our real-time PCR assays, we first extracted poly-A mRNA from selected tissues with the Dynabeads kit from DYNAL. Based on poly (dT) magnetic beads, this kit allowed us to isolate poly-A mRNA of good quality from various tissues. The *L.sanguineus* poly-A mRNA samples were then immediately reverse transcribed by the Superscript III kit from Invitrogen using random hexamers. The use of oligo (dT) for reverse transcription may lead to incomplete 5'end cDNA synthesis of some long RNAs and/ or RNAs exhibiting strong secondary structures. As this would lead to an under-representation of some mRNAs in our samples, and consequently bias the interpretation of our expression data, we exclusively used random hexamers for the reverse transcription step.

The primers used in our assays were designed with a specific program available from the website www.genscript.com/ssl-bin/app/primer. Primers were designed in a way to produce amplicon excluding the homeobox in order to reduce possible mispriming. In accordance with the recommendations of the real-time PCR kit supplier, the size of every amplicon has to be between 100 and 150 bp. Subsequently, the primers were tested in a control run without cDNA to make sure that primers did not dimerize. Several combinations were tested on a mixture of cDNA obtained from mRNA extracted from diverse body region origin. We selected the pair of primers performing the best results in our control experiment. Especially, we controlled the absence of primer dimerization, the reproducibility of the run and the numbers of runs needed to reach the threshold of detection of the real-time PCR

product by the machine. An optimization of the reverse transcription step was also required. We made sure we were always starting the experiment with the same range of mRNA concentration and that we were using the same incubation time. Furthermore, for each experiment, at least two independent mRNA extractions and subsequent reverse transcriptions were performed. In addition, three technical replicates of real-time PCR experiments were carried out. This procedure allows us to get statistically significant values and to calculate their standard deviation values.

3.3.2 Analysis *Ls-Otx* and *Ls-Cdx* expression levels by real-time PCR in adult *L.sanguineus*

As a further control of our real-time PCR experimental procedures (mRNA quality, cDNA synthesis efficiency, primers couples, cycle numbers…), we decided to look for *Ls-Otx* and *Ls-Cdx* expression levels in various adult body parts. We prepared cDNA from mRNA extracted from three different body regions: the whole cerebral region (region 1 to 3), the gut (from region 4 to 8) and the most posterior body intestine part, plus the anus (from region 8 to 10). By comparing the data obtained by real-time PCR with the previous ones from ISH, we should be able to validate our real-time PCR experimental procedures or to disprove them.

The analysis of *Ls-Otx* expression levels by real-time PCR (Fig. 2.14) nicely corroborates our ISH data (Fig. 2.11) and shows that *Ls-Otx* is expressed in the head part of the worm. We could detect expression of *Ls-Cdx* in the adult gut (Fig. 2.14), while no staining was detectable by ISH. However, differences between ISH data and real-time PCR data for *Ls-Cdx* can be explained in two possible ways: either the weak signal observed by ISH in the intestine of some adults is interpreted as background and therefore, the expression level of *Ls-Cdx* in adults is too low for detection by ISH methods. Or, in contrast to our initial interpretation of the *Ls-Cdx* ISH data, we can assume that the weak ISH signal observed in adults actually reflects some weak *Ls-Cdx* expression.

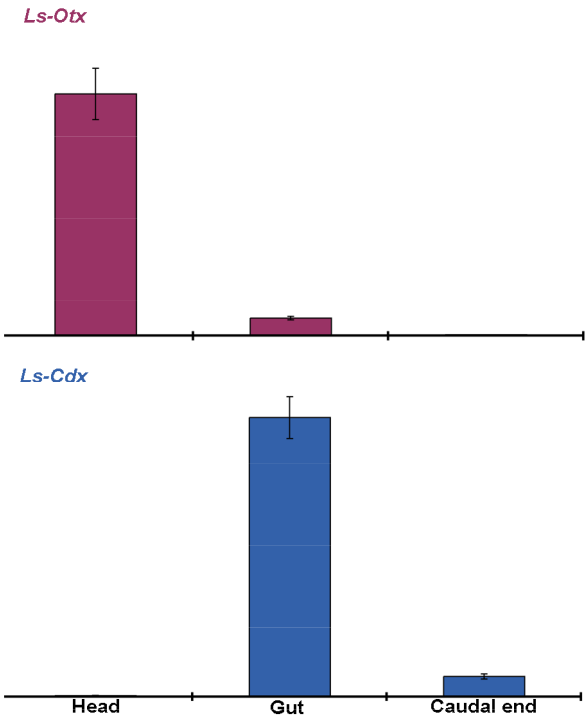


Fig. 2.14 *Ls-Otx* and *Ls-Cdx* expression levels measured by real-time PCR of different *L.sanguineus* body parts
The graphs display relative values normalized to elongation factor (EF1α) expression level. cDNA was synthesized from different mRNA preparations. cDNA synthesized from mRNA extracted from the whole cerebral (region 1 to region 3) are referred as "head". The ones from the gut, including esophagus and intestine (region 4 to middle of region 8), are referred as "gut". The ones from the most posterior intestine part and anus regions referred as "caudal end".

3.4 Analysis of *Ls-Otx* and *Ls-Cdx* expression during both, anterior and posterior regeneration of an isolated fragment 8

We have shown that *Ls-Otx* and *Ls-Cdx* are activated during brain regeneration (Fig. 2.12) and gut regeneration (Fig. 2.13), respectively. Interestingly, the expression pattern of *Ls-Cdx* is asymmetric during intestinal regeneration: we showed that *Ls-Cdx* is only expressed at one extremity of an isolated fragment 8 that undergoes both, anterior and posterior regeneration (Fig. 2.13, B). These isolated fragments, from a single adult worm, undergo regeneration at both, their anterior and posterior ends. This leads to the production of clonal somatic embryos. We decided to analyze more precisely the expression of *Ls-Otx* and *Ls-Cdx* during the regenerative processes occurring at both extremities of these isolated fragments 8, by ISH and by real-time PCR approaches.

3.4.1 ISH approach:

Reasonably large adult *L.sanguineus* were chosen for the experiment. They were starved for two weeks before starting the experiment to avoid background problem due to digestive enzyme activity. Multiple transections within the long intestinal-genital module (region 8) of a single worm were performed; isolated fragments were subsequently kept in the dark at 18°C. They were then fixed sequentially after 3, 4, 6, 8, 10, 15 and 22 days. Several replicates were prepared for each time point. In agreement with the experimental procedures described above, all regenerating fragments, presented in figure 2.15, originate from the same worm. Hence, the "pre-regenerating conditions" of every specimen, such as the age of the organism and its physiological status, are identical.

Both, *Ls-Otx* and *Ls-Cdx* are activated in regenerating intestinal-genital fragments (Fig. 2.15). *Ls-Cdx* exhibits a clearly asymmetric activation. It is expressed at only one extremity throughout the entire process of regeneration. In contrast, *Ls-Otx* is activated on both regenerating sides in the earliest stages (Fig. 2.15, A1 to A3). Later, its expression becomes restricted to only one extremity (Fig. 2.15, A4 to A7).

Ls-Otx is expressed at both extremities of the regenerating intestine fragment for approximately the first week post transection. Initially, during the first four days, its expression level seems to be identical on both sides. However, mRNA transcript abundance is progressively declining on one side (Fig. 2.15, A3): *Ls-Otx* expression is finally restricted to one extremity, only (Fig. 2.15, A4 to A7). This extremity corresponds to the future anterior end of the somatic embryo, judging from the blastema morphology. From the blastema shape of this stage, we are indeed able to distinguish the anterior from the posterior end of the "somatic embryo": the anterior part forms a prominent "bud", from which cerebral organs will soon start to differentiate, whereas the regenerative process is delayed at the opposite posterior extremity and only an epithelial invagination can be observed. After two weeks (Fig. 2.15, A6), we observe *Ls-Otx* expression in the regenerating cerebral organs and their canals. It is also strongly expressed in the regenerating cerebral ganglia.

After three weeks (Fig. 2.15, A7), a small entire worm has regenerated. *Ls-Otx* is still expressed in the head of this worm. A more detailed observation of its expression pattern is in good correlation with the expression data already described for an adult (Fig. 2.11).

Ls-Cdx is always expressed only at one side of the somatic embryo. In the early stages (Fig. 2.15, B1 to B4), we observe a diffuse staining, at the level of epithelial invaginations. This expression domain is less broad ten days after the transection (Fig. 2.15, B5). After two weeks, the *Ls-Cdx* expression level is weaker and even more restricted to the extreme posterior tip of the developing embryo (Fig. 2.15, B6). Consistent with the ISH data obtained for adults, *Ls-Cdx* is not detectable anymore in the fully regenerated worm (Fig. 2.15, B7).

From day fifteen post-transection (Fig. 2.15, A6 and B6), the A-P axis is clearly distinguishable, based on blastema morphology. From earlier time-points, the anterior regenerating extremity and the posterior one are difficult to distinguish. Nevertheless, we assume that the *Ls-Otx* expression we detect during this period was present at the anterior extremity, while *Ls-Cdx* was expressed at the posterior one. To make sure that these two genes are expressed at opposite sides, we performed a double ISH to detect both mRNA in the same sample (Fig. 2.16). We found that the *Ls-Otx* and *Ls-Cdx* expression domains are indeed non-overlapping in 8 days post-transection (dpt) regenerating intestinal fragments (corresponding to A4 and B4 stages described in Fig. 2.15): *Ls-Cdx* is activated at the opposite side of *Ls-Otx*. Although we can't discriminate both extremities based on anatomical characteristics at these stages, *Ls-Otx* and *Ls-Cdx* are very likely expressed at the future anterior and the future posterior end, respectively.

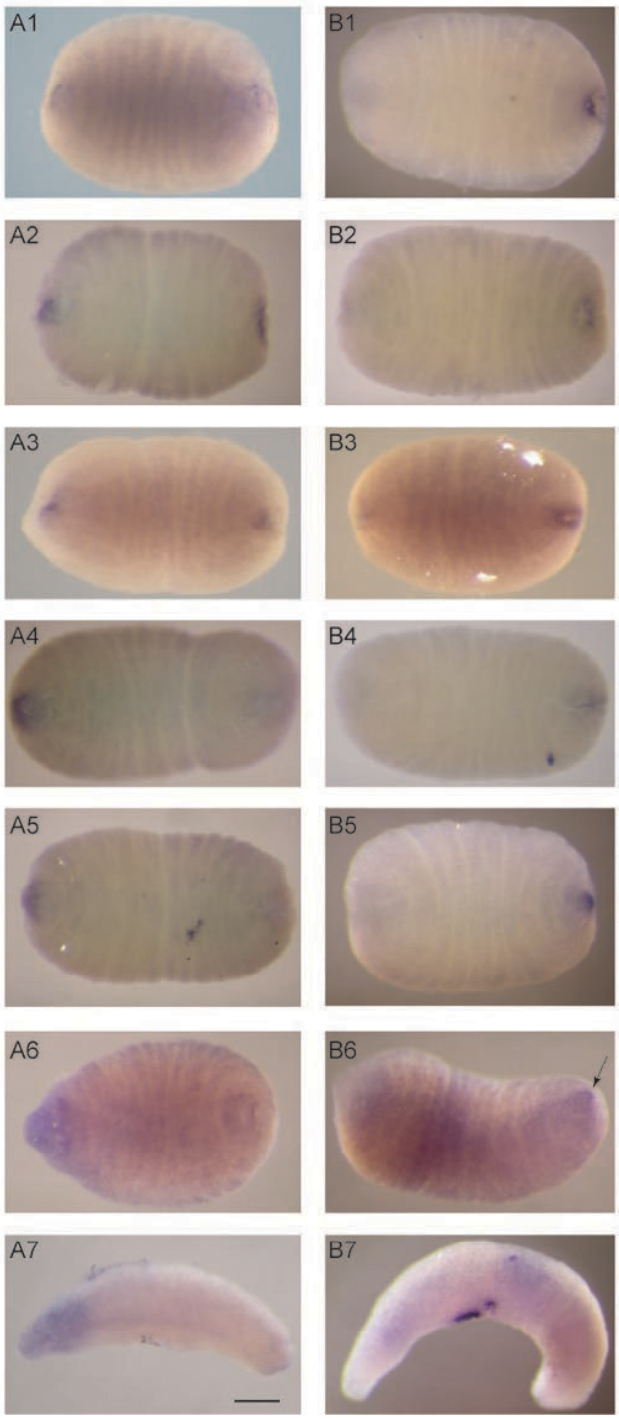


Fig. 2.15 *Ls-Otx* and *Ls-Cdx* expression in isolated body fragment 8 undergoing both anterior and posterior regeneration

When an A-P polarity is distinguishable, based on the morphology of the blastemata (from A/B4 to A/B7), samples have been oriented with anterior to the left. Ventral view for all, except for A7 and B7, from which their left side are observed.

A1 to A7: ISH were done using a *Ls-Otx* anti-sense RNA probe. B1 to B7: ISH were done using a *Ls-Cdx* anti-sense RNA probe.

A1, B1: 3 dpt (day post transection) / A2, B2: 4 dpt / A3, B3: 6 dpt / A4, B4: 8 dpt / A5, B5: 10 dpt / A6, B6: 15 dpt / A7, B7: 22 dpt.

Arrow in B6 points at the weak *Ls-Cdx* expression localizing to the posterior tip of the specimen. Scale bar: for A1 to A6 and B1 to B6: 300 μm / for A7 and B7: 500 μm

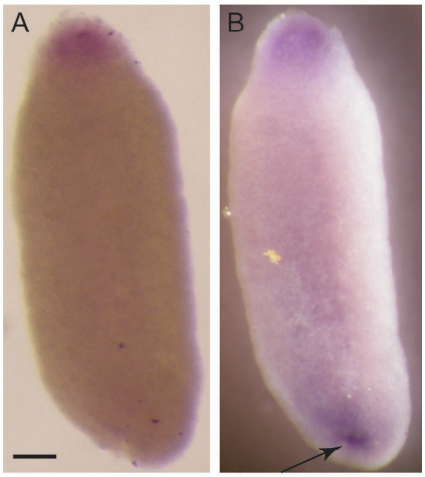


Fig. 2.16 Double ISH with *Ls-Otx* and *Ls-Cdx* anti-sense RNA probes on a regenerating isolated intestinal fragment, 8 days post-transection

A FITC-labeled anti-sense RNA probe was used for *Ls-Otx* detection, a Digoxygenin-labeled anti-sense RNA probe for *Ls-Cdx* detection. Scale bar: 250 μm.

A: first color detection of *Ls-Otx* expression with Magenta phosphate substrate. Its reaction with the AP coupled to the anti-FITC Ab gives a purple precipcitate

B: second color detection on the same sample than in A of *Ls-Cdx* expression with NBT/ BCIP substrate. Its reaction with the AP coupled to the anti-Dig Ab gives a blue precipitate.

The arrow points at *Ls-Cdx* expression, at the opposite extremity of *Ls-Otx* expression.

The purple staining for *Ls-Otx* expression in A seems to get a bit blue in B. This is due to a technical artefact and not to an expression of *Ls-Cdx* at both extremities. We know from previous ISH data that *Ls-Cdx* is only expressed at one side, the posterior end, at this stage of regeneration.

While undergoing regeneration, *L.sanguineus* tissues produce more mucus than normally. The presence of such mucus is a major technical problem in ISH experiment. We first tried to remove as much mucus as possible with forceps. Then, we applied a cystein chloride treatment, after a quick pre-fixation of tissue with paraformaldehyde. This procedure was not always sufficient to ensure complete removal of mucus causing background problems, especially in the earliest stages of regeneration.

We could not obtain decent ISH data for regenerating intestine fragments, prior to 72 hours post-transection. Therefore, in order to investigate whether *Ls-Otx* and *Ls-Cdx* are expressed at these earliest regenerating stages, we decided to use a real-time PCR approach.

3.4.2 Real-time PCR approach

In order to investigate the expression of *Ls-Otx* and *Ls-Cdx* in early regenerating intestinal fragment, prior to 72 hours post-transection, we used a real-time PCR approach. Isolated intestinal fragments, undergoing regeneration, were quickly frozen in liquid nitrogen at different time-points: 0, 1, 2, 3, 5, 6, 8, 9, 10, 11, 12, 13 and 14 days post-transection. They were further kept at -70°C. In order to control the possible variations of quality in the procedure, mRNA extraction and subsequent reverse transcription were done at the same time for all samples of one experiment run.

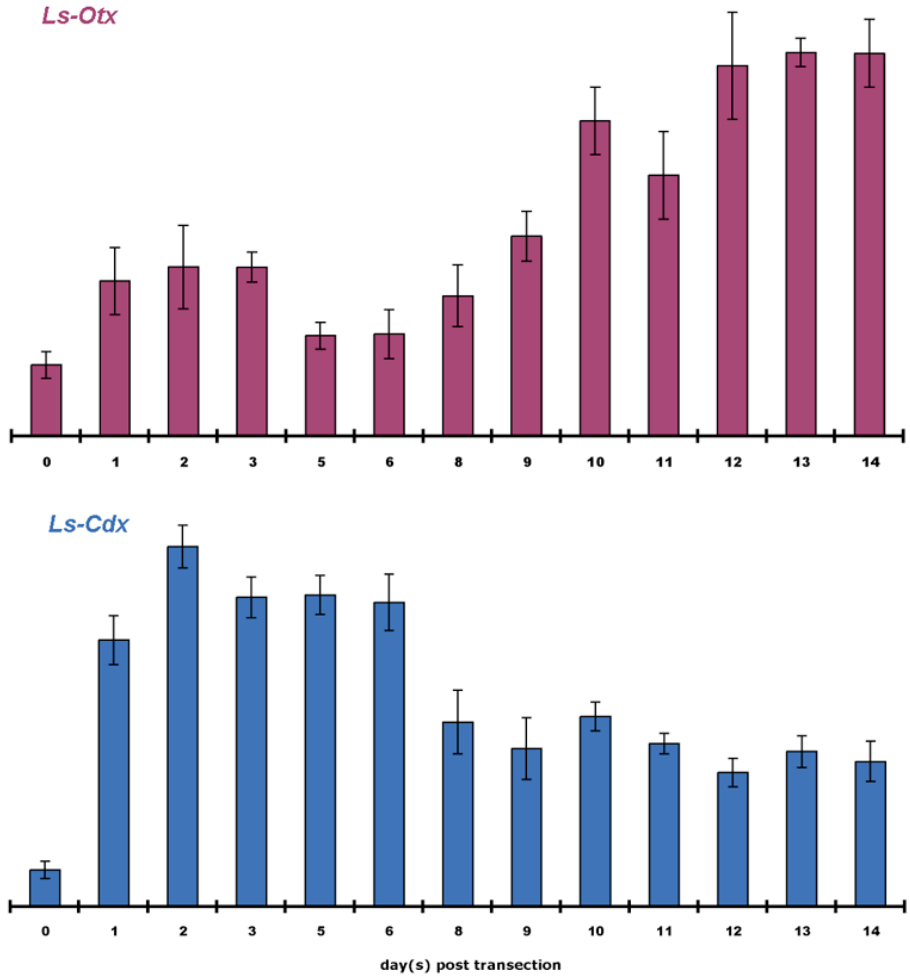


Fig. 2.17 Expression levels of *Ls-Otx* and *Ls-Cdx* during both anterior and posterior regeneration of isolated intestinal fragment deduced by real-time PCR
Ls-Otx and *Ls-Cdx* expression levels are determined by real-time PCR. They are normalized relative to the level of expression of EF1α. As a control, the expression levels of *Ls-Otx* and *Ls-Cdx* have been determined in an isolated intestinal fragment, immediately after transection. This control is referred as the sample "0 day post amputation".

The results obtained by ISH (Fig. 2.15) have been confirmed by the real-time PCR analysis: *Ls-Otx* and *Ls-Cdx* are both up-regulated during bi-directional regeneration of intestinal fragment (Fig. 2.17). However, their activations present different kinetic profiles. Both are activated within the first 24 hours. However, *Ls-Cdx* already reaches its maximum values within the first days, whereas *Ls-Otx* expression levels become highest after ten days of regeneration (Fig. 2.17).

Ls-Otx activation seems to occur in two phases (Fig. 2.17). A rapid up-regulation of *Ls-Otx* occurs during the first three days post transection, followed by a weak decrease of its expression level. Then, one-week post transection, *Ls-Otx* is constantly up-regulated, reaching a plateau. This expression level probably corresponds to its normal expression level in a fully-grown head. From our ISH results, we can postulate that *Ls-Otx* is expressed during two different phases of regeneration: in earliest regenerative stages, it is expressed at both extremities. Later, *Ls-Otx* expression becomes restricted to the extremity, which undergoes anterior regeneration only.

Ls-Cdx is strongly up-regulated in the days following the transection. In comparison to their respective normal levels of expression in a non-regenerating intestine, *Ls-Cdx* is much stronger up-regulated than *Ls-Otx* during the first week post transection (Fig. 2.17, samples 1 to 8). The maximal level of *Ls-Cdx* expression is reached within the first two days post transection, while the maximal level of *Ls-Otx* expression is reached only ten days later. *Ls-Cdx* stays strongly activated for approximately one week and then, starts to be down-regulated. This down-regulation of *Ls-Cdx*, after one week, corroborates our results obtained from ISH experiments: one week after the transection, in comparison with earlier regenerating stages, we detected a weaker *Ls-Cdx* expression, which is restricted to the posterior tip of the regenerating fragment (see Fig. 2.15, B6).

3.5 Analysis of *Ls-Otx* and *Ls-Cdx* expression during anterior and posterior regeneration of worms missing body regions 1 to 5 and worms missing body regions 5 to 10, respectively.

We have assumed that, during the regeneration experiment involving isolated intestinal fragment (paragraph), both, anterior and posterior regeneration were occurring to give rise to an entire new worm. Anterior regeneration has certainly to occur in order to achieve the replacement of all anterior missing regions (from region 1 to region 7). However, adults *L.sanguineus* are often lacking an anus. Therefore, whether the posterior end is fully regenerating is less certain. Thus, the process, which takes place at the posterior end, could be more related to a wound healing process rather than a posterior regeneration process.

Hence, to investigate *Ls-Cdx* expression during posterior regeneration, we decided to perform the same analysis by ISH and real-time PCR on regenerating body parts, from other origins. Using a razor blade, adults *L.sanguineus* were cut into two parts: as the cut was performed behind the mouth, we obtained worm pieces composed of region 1 to 5: “upper parts” and worm pieces composed of region 6 to 10: “bottom parts” (Fig. 2.18). Theoretically, we know that the bottom parts will regrow a head and the upper ones, a tail. Hence, on one hand, the “bottom parts” allow us to analyze once more anterior regeneration process. On the other hand, as *L.sanguineus* need to regenerate their gut in order to survive, we have the opportunity to analyze a true posterior regeneration from the “upper parts”.

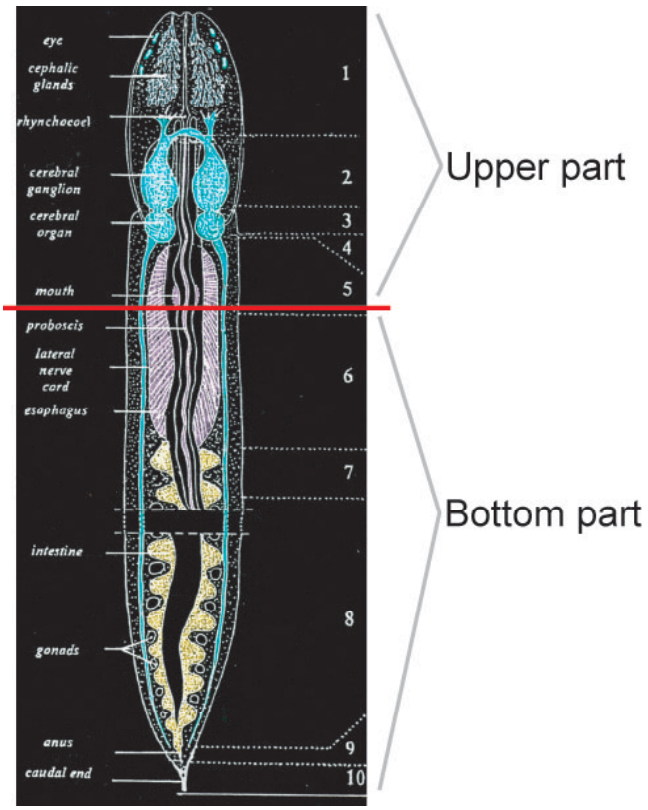


Fig. 2.18 Adult *L.sanguineus* amputation scheme
A red line is present at the level of the amputation plane. This amputation results in the isolation of two fragments of the worm: an "upper part" and a "bottom part".
The "upper part", composed of body regions 1 to 5, is lacking the body regions 6 to 10. The "bottom part", composed of body regions 6 to 10, is lacking the body regions 1 to 5. In order to restore the normal *L.sanguineus* body pattern, the "upper part" and the "bottom part" undergo posterior and anterior regeneration, respectively.

3.5.1 ISH approach

We wanted to investigate different stages of anterior and posterior regeneration occurring in "bottom parts" and "upper parts", respectively. Hence, we performed razor blade amputations 16, 12, 6 and 4 days before the fixation of the samples with paraformaldehyde. Technically, we used at least two worms per time-point and we performed the paraformaldehyde fixation of all body fragments on the same day.

3.5.1.1 *Ls-Otx* and *Ls-Cdx* expression during anterior regeneration from worms missing body regions 1 to 5: " bottom parts"

Whereas *Ls-Cdx* is not expressed during anterior regeneration (Fig. 2.19, C1 to C4), *Ls-Otx* expression is observed in all the regenerative fragments. By analyzing bigger worm fragments in this experiment, we can describe more precisely the expression pattern of *Ls-Otx* during anterior regeneration.

Following the amputation, an intense constriction of the ring-shaped muscular tissues reduces quickly the size of the opened wound at the anterior extremity of the worm. Rapid extension of the nearby epidermis seals the wound and covers the injured tissues by forming a ciliary and secretory healing epithelium. The discoid ciliated (without secretory cells) epithelial region, named DCER, is made up of a cylindrical epithelium, exclusively composed of closely joined ciliated columnar cells (Fig. 2.19, A and B). The adjacent external epithelium contains the three usual epithelial components: columnar ciliary cells mixed with both granular mucous and homogeneous serous cells (Fig. 2.19, C). The DCER differentiates and increases in size from the central region of the initial wound epithelium. This involves, first, disruption of the columnar arrangement of the epithelial cells and, secondarily, apoptosis of all the secretory cells. Thus, an annular region located between the newly DCER and the yet-unmodified peripheral secretory epithelium, contains all the cellular types of the epidermis, mixed together (G. Rué and D. Brossard, personal communication). *Ls-Otx* is specifically detected in this annular intermediate region. At the level, of *Ls-Otx* expression, the mucous and serous glandular cells locate deeper in the perturbed epithelial structure and become elongated. Their new, stretched morphology correlates with their translocation towards the underlying muscular and conjunctive tissues, where they enter apoptosis. Important processes, such as epithelial disorganization, cell differentiation and apoptosis program entry occur at the level of the annular region between the DCER and the adjacent external epithelium (Fig. 2.20, A and B). As *Otx*, is strongly up-regulated at this level, it may play a pivotal role in those processes. Furthermore, we can clearly distinguish two patches of cells strongly expressing *Ls-Otx* at the amputation plane.

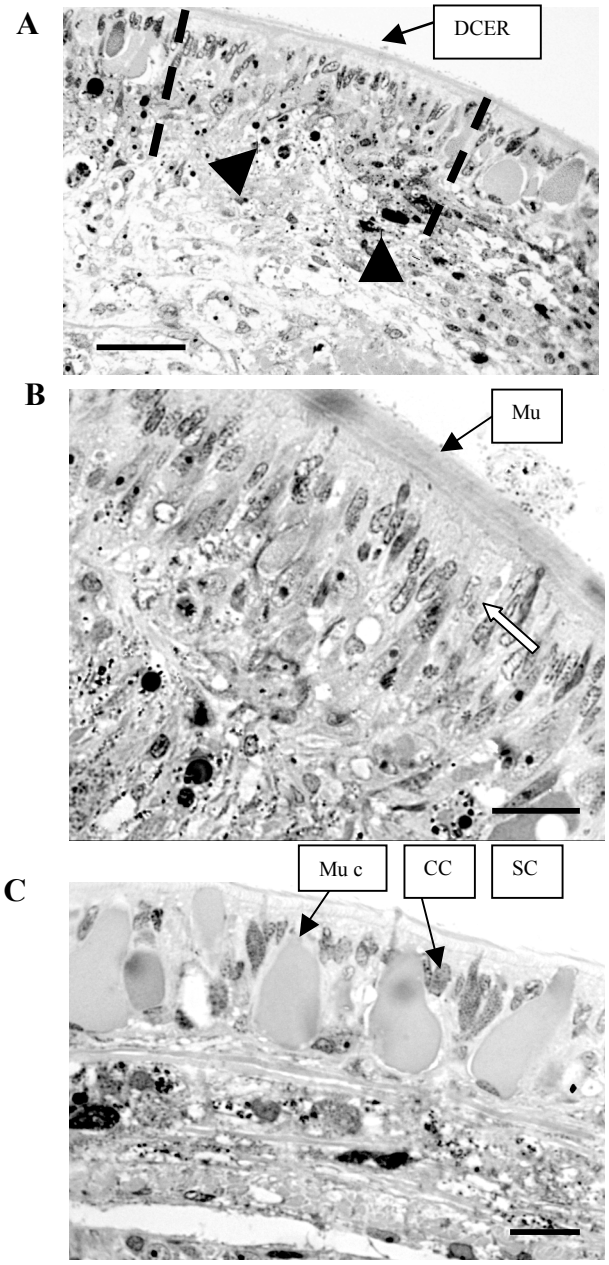


Fig. 2.19 One-week-old *L.sanguineus* somatic embryo: details of the histological structure of the healing regions (From G.Rué and D.Brossard).
A: Transversal section of the thin discoid ciliated epithelial region (DCER). It contains no secretory cells. The DCER enlarges progressively from the cicatricial glandular/ciliated epithelium (the dotted lines correspond to the peripheral limits of the DCER). Disorganized tissues made with necrotic or apoptotic cells lie just below the DCER (arrowheads). Scale bar: 10 μm.
B: Detail of the DCER, which is characterized by the regular alignment of the nuclei of the ciliated cells (white arrow). A homogeneous mucous deposit covers the epithelial ciliated cells. Scale bar: 5 μm.
C: Histological structure of the adjacent secretory epithelium made with mucous (Mu C), serous (SC) and ciliated (CC) cells. Scale bar: 5 μm.

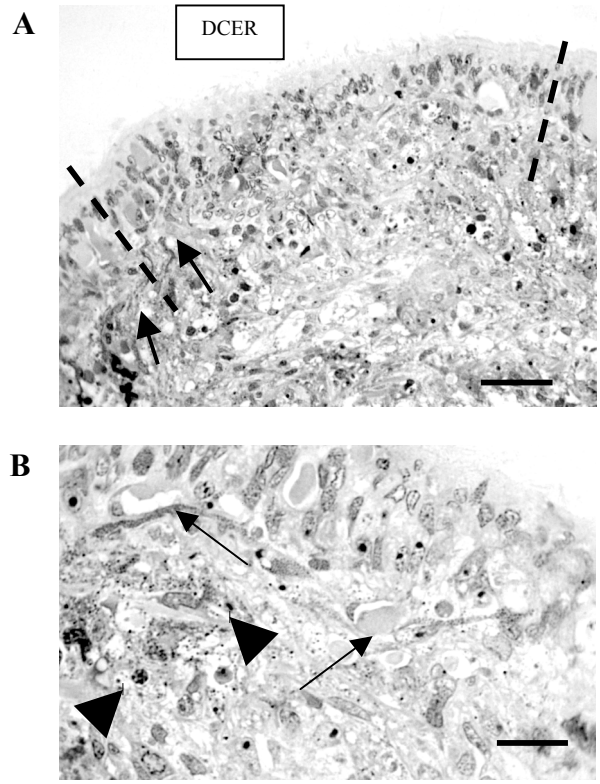


Fig. 2.20 Differentiation of the DCER during somatic embryonic development (From G.Rué and D.Brossard).
A The DCER differentiates in a region where the usual organization of the epithelial cells is totally disrupted. In contrast to the regular alignment of the mixed glandular and ciliated cells present in the usual epidermic epithelium (Fig. 2.19, C), the secretory serous and the mucous cells (arrows) that are located at the limits of the DCER (doted lines) elongate towards the underlying tissues where a lot of them appear apoptotic/necrotic. Scale bar: 5 μm.
B. Enlargement of the region located at the limits of the DCER and the adjacent typical epidermic epithelium. The characteristic alignment of the epithelial cells is disturbed and secretory cells (arrows) appear elongated in the underlying tissues, in close proximity with apoptotic/necrotic areas (arrowheads). Scale bar: 10 μm.

Those patches correspond to the tip of the lateral nerve cords, which have been sectioned by the amputation. The cells of the regenerative head bud locate between the DCER and the proximal "cytolytic area", which contains a lot of granules-rich phagocytic cells and probably corresponds to loci, where histolytic products accumulate. As mitoses are rarely observed in this region during that period, it has been suggested that these regenerative cells originate from different, not-well established, loci of the worm fragment. They probably immigrate from the parenchyma or the proximal regions of the epidermis (see Gibson, 1972). Some histological features, such as the large epithelial invaginations that develop close to

the periphery of the anterior DCER, are characteristic of early regenerative stages. These large invaginations appear deeply embedded in the inner epidermis and consist of ciliary, secretory, basal and mitotic cells. Selected epithelial cells rapidly translocate from the surface, through the large epithelial invaginations, to the inner, actively reorganized, anterior blastema region. (G. Rué and D. Brossard, personal communication). Interestingly, as shown in B2 from Fig. 2.21, *Ls-Otx* is expressed at the level of these large invaginations. It is also detectable at the tip of the sectioned lateral nerve cords (Fig. 2.21, B1 and B2). Moreover, we detect a broader *Ls-Otx* expression at the level of the dense anterior regenerating sub-epithelial tissues (Fig. 2.21, B2 and B3). Due to collective transfers of different epithelial cells, mostly through the large invaginations, these regenerating sub-epithelial tissues become rapidly larger, before the cerebral organs and other organs differentiate in the cephalic area. The most anterior module lacking (region 1) is the first to organize during somatic embryogenesis, nevertheless the time to obtain complete differentiation of the antecerebral module is longer than the time necessary to differentiate the cerebral ganglion module (region 2) and the cerebral organ module (region 3). Differentiation of cerebral ganglia, as numerous cellular bodies and cerebral organs as paired voluminous cellular masses become very obvious. *Ls-Otx* is expressed in this differentiation process (Fig. 2.21, B3 and B4) and therefore it might to be involved in neurogenesis occurring during somatic embryogenesis and also in the differentiation of sense organs.

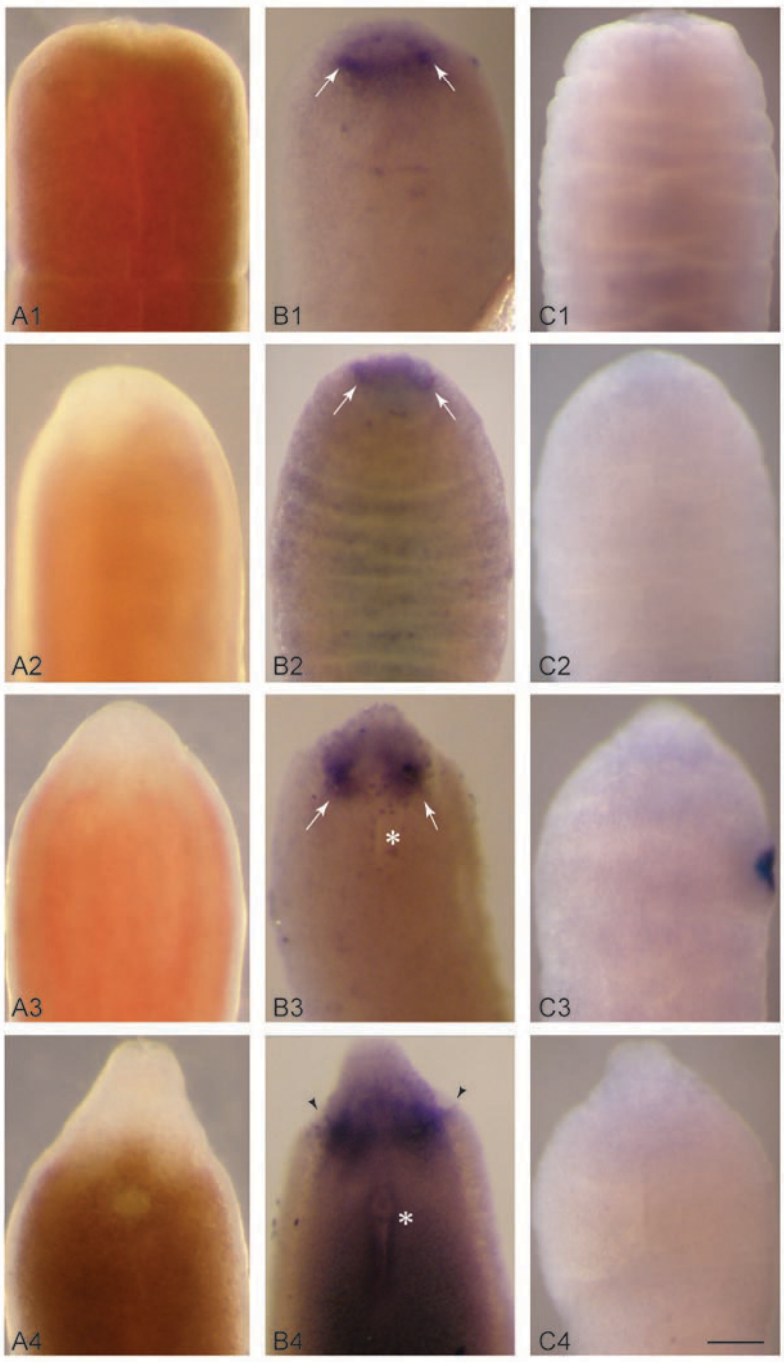


Fig. 2.21 *Ls-Otx* and *Ls-Cdx* expression during anterior regeneration from *L.sanguineus* lacking body regions 1 to 5

ISH were done using a *Ls-Otx* anti-sense RNA probe in B and a *Ls-Cdx* anti-sense RNA probe in C.

A: pictures taken before tissues fixation

A1, B1, C1: 4 dpa (day post amputation) / A2, B2, C2: 6 dpa / A3, B3, C3: 12 dpa / A4, B4, C4: 16 dpa

Anterior is oriented to the top. Scale bar: 1 mm

White arrows indicate *Ls-Otx* expression in lateral nerve cords at the amputation plane in B1 and epithelial invaginations in B2. Note that regeneration of ventral cerebral ganglia occurs in continuity with the existing lateral nerve cords (white arrows in B3). A new mouth develops rapidly (asterisk). Invaginations forming the cerebral organs canals are indicated by arrowheads in B4.

Due to the size of the samples used in this experiment, we were able to describe more precisely the expression pattern of *Ls-Otx* during anterior regeneration, than we did in the precedent experiment. The regeneration of anterior structures occurs more rapidly in worms amputated behind the mouth than in isolated intestinal fragments. However, the anterior regenerative processes that we described in this paragraph also apply to the ones, which take place in an isolated intestinal fragment.

In contrast to *Ls-Otx*, *Ls-Cdx* is not expressed during the anterior regeneration process, described here (Fig. 2.21, C1 to C4). This is in accordance with the fact that, in an isolated intestinal fragment, it is only expressed at the end, which undergoes a posterior regeneration (Fig. 2.15 and Fig. 2.16).

3.5.1.2 *Ls-Otx* and *Ls-Cdx* expression during posterior regeneration from worms missing body regions 6 to 10: "upper parts".

In order to investigate the expression of *Ls-Otx* and *Ls-Cdx* during posterior regeneration, we searched for their expression in the "upper parts", left from the razor blade cut done behind the mouth of the animal. These "upper parts" are composed of body regions 1 to 5. Although posterior regeneration begins later and is slower than anterior regeneration, it depends on the same primary processes, such as rapid wound healing, followed by the differentiation of a discoid ciliated epithelial region and epithelial invaginations. The epithelial invaginations, which are located at the posterior extremity of the regenerating worm, appear to be tighter than the ones, which develop during an anterior regeneration (G.Rué and D.Brossard, personal communication). In addition, posterior blastemata are smaller and thinner than anterior ones.

The restoration of the appropriate proportion and pattern in *L.sanguineus* is thought to involve remodeling of existing tissues. For example, when undergoing posterior regeneration of body regions 6 to 10, *L.sanguineus* worms get thinner while they elongate their A-P body axis (Fig. 2.22). Interestingly, upon prolonged periods of starvation, these worms also shrink in body mass and can be reduced in length to only few millimeters.

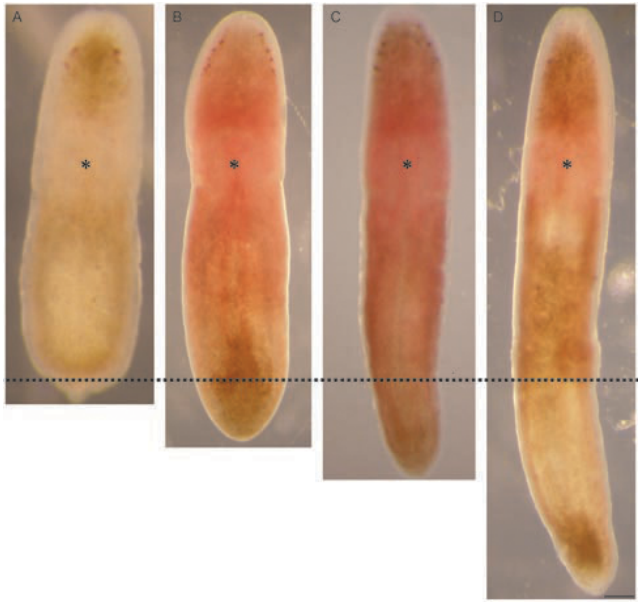


Fig. 2.22 Posterior regeneration from adult *L.sanguineus* lacking body regions 6 to 10 is epimorphic
Pictures were taken before fixation; worms were anesthetized with 8% magnesium chloride. We look at worms from their dorsal side, with anterior orientated to the top.
A: 4 dpa / B: 6 dpa / C: 12 dpa / D: 16 dpa
A dashed line is present at the level of the amputation plane. Asterisks indicate the position of the brain.
Scale bar: 650 μm

Ls-Cdx is specifically expressed during posterior regeneration (Fig. 2.23, B1 to B4). Four days after the amputation, we can observe that *Ls-Cdx* is expressed at the posterior end of the regenerating worm, probably in the gut (Fig. 2.23, B1). Later, at six and twelve days post amputation (Fig. 2.23, B2 and B3), its expression becomes restricted to the extreme posterior tip of the worm. This corresponds to expression in the small developing posterior blastema. After the onset of posterior regeneration, *Ls-Cdx* is de-novo expressed more broadly, in close vicinity with the discoid ciliated epithelial region, in the posterior regenerating sub-epithelial tissues (Fig. 2.23, B3 and B4).

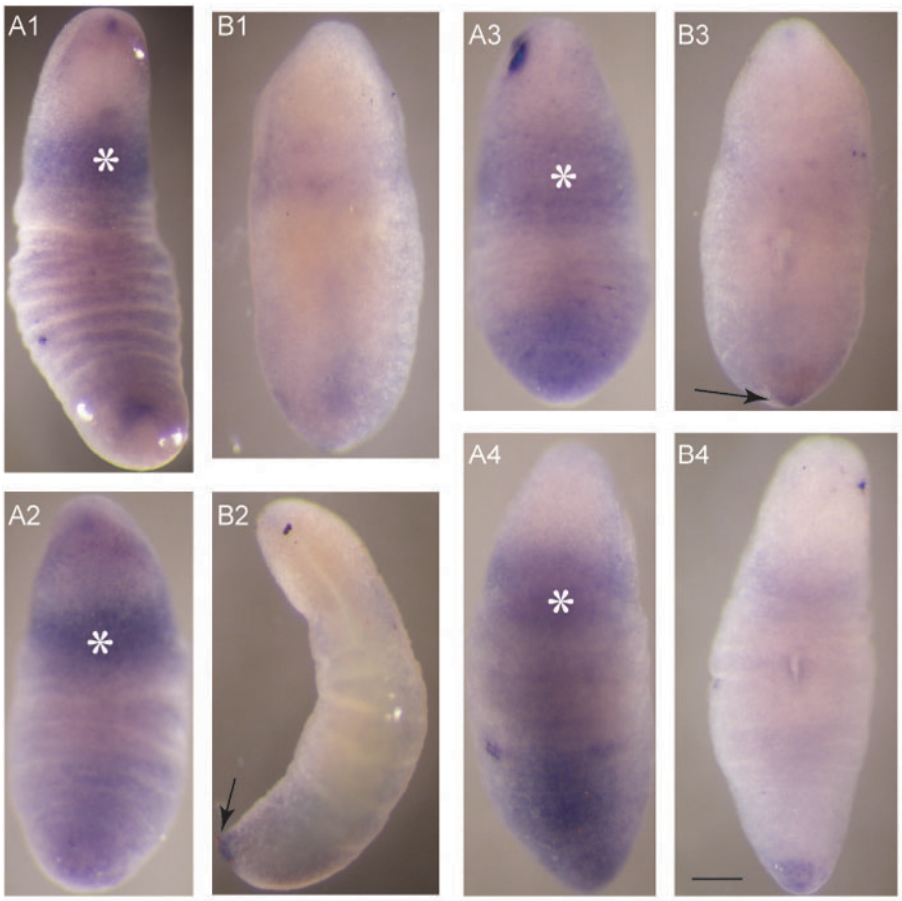


Fig. 2.23 *Ls-Otx* and *Ls-Cdx* expression during posterior regeneration from *L.sanguineus* lacking body regions 6 to 10
ISH was done by using a *Ls-Otx* anti-sense RNA probe in A and a *Ls-Cdx* anti-sense RNA probe in B.
A1, B1: 4 dpa / A2, B2: 6 dpa / A3, B3: 12 dpa / A4, B4: 16 dpa
Anterior is oriented to the top. Animal observed from its left side in B2. Scale bar: 1 mm.
Asterisks indicate a similar *Ls-Otx* expression in brain and cerebral organs, as the one usually observed in non-regenerating adult.
Note the presence of background at the level of esophageal modules, especially in A1 to A4 samples.
Arrows indicate precise *Ls-Cdx* expression pattern at the extreme posterior regenerating tip in B2 and B3.

Even though we tried several experimental conditions for the ISH protocol, we didn't obtain clear *Ls-Otx* expression data during posterior regeneration of the "upper parts" (Fig. 2.23, A1 to B4). We can observe a similar *Ls-Otx* expression in the cerebral ganglia and the sense organs, as the one observed in non-regenerating adults (Fig. 2.11). However, we cannot surely discriminate staining from background in the regenerative posterior ends of those organisms.

3.5.2 Real-time PCR approach:

We didn't obtain clearly interpretable ISH data for *Ls-Otx* expression during posterior regeneration of a worm missing body regions 6 to 10. In addition, due to an abundant production of mucus in the earliest regenerating stages, it was not possible to get reproducible *Ls-Cdx* expression pattern for these stages. Therefore, in order to investigate the expression of *Ls-Otx* and *Ls-Cdx* in the earliest posterior regenerating stages, we decided to use a real-time PCR approach.

We amputated the worms on different time-points in order to get samples, undergoing regeneration for 13, 11, 9, 5, 3, 2 and 1 days, at the time we performed the mRNA extraction and the subsequent reverse transcription of all "upper parts". At least, two worms were used per time point. The second body fragment, the "bottom part", which results from the amputation, was kept to maintain our laboratory *L.sanguineus* stock.

Results from the real-time PCR experiment confirm the up-regulation of *Ls-Cdx* during posterior regeneration of the body regions 6 to 10 of the worm fragment, composed of body regions 1 to 5 (Fig. 2.24). It should be noted that no *Ls-Cdx* mRNA is detected in the non-regenerating fragment, used a control (Fig. 2.24, sample "0"). In comparison with its level of expression in the control sample, *Ls-Cdx* is already activated within 24 hours and reaches its maximum values within the first week of regeneration.

In comparison with its level of expression in a non-regenerating fragment (Fig. 2.24, sample "0"), *Ls-Otx* is up-regulated within 24 hours. However, this up-regulation seems to be of short duration: 2 days after the amputation, the *Ls-Otx* level of expression is back to the level measured in the non-regenerating fragment, used a control. This level of expression corresponds to the normal expression of *Ls-Otx* in the adult brain and in the adult sense organs (see Fig. 2.14 for the level of *Ls-Otx* expression in an adult *L.sanguineus* head, determined by real-time PCR).

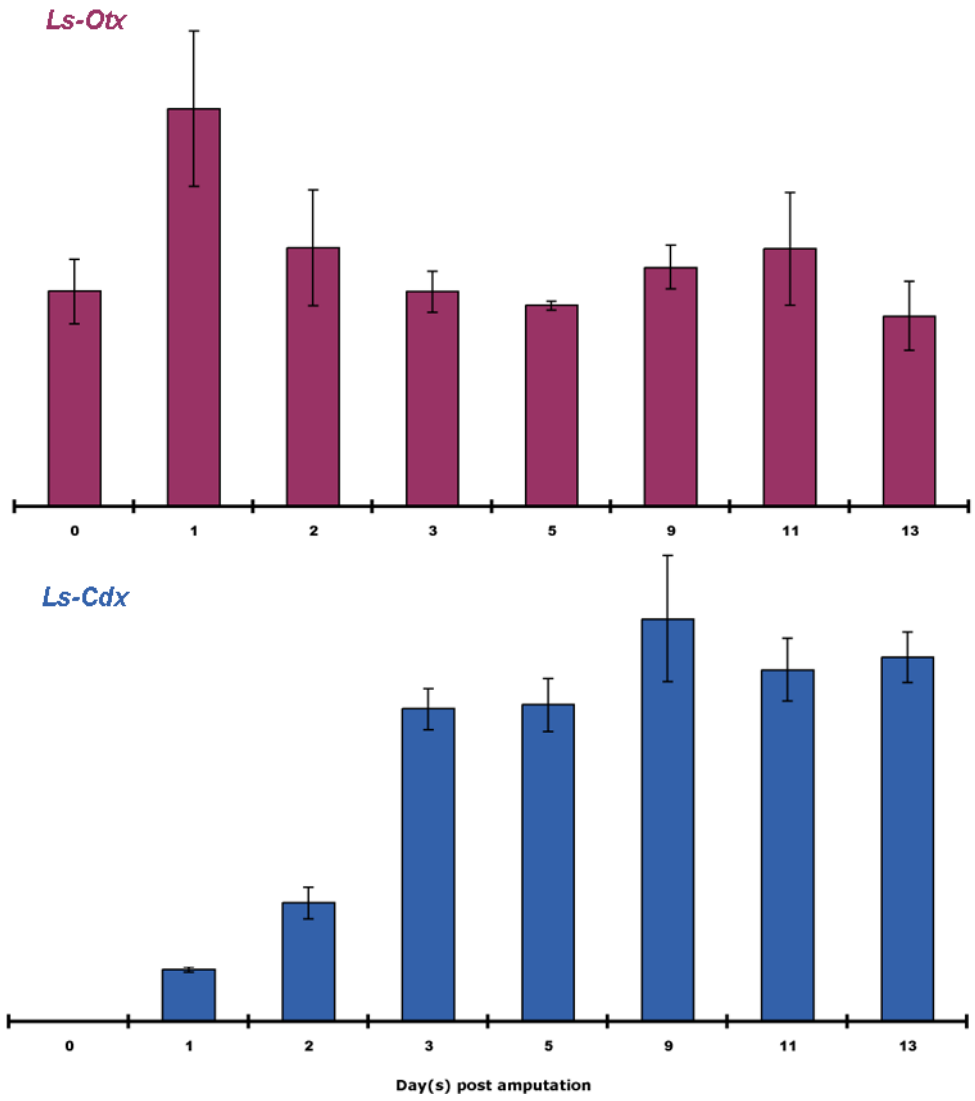


Fig. 2.24 real-time PCR expression of *Ls-Otx* and *Ls-Cdx* during posterior regeneration of *L.sanguineus* amputated from body regions 6 to 10

Ls-Otx and *Ls-Cdx* levels of expression are evaluated by real-time PCR and presented relative to the normalizing value of elongation factor. As a control the expression levels of *Ls-Otx* and *Ls-Cdx* have been determined in an "upper part" (body fragment composed of regions 1 to 5), fixed in paraformaldehyde immediately after the amputation. This control is referred as the sample "0 day post amputation".

3.6 Investigation of *Otx* expression in *L.lacteus*, a close relative of *L.sanguineus*, which exhibits different regeneration properties.

By ISH (Fig. 2.15, 2.16 and 2.19) and real-time PCR experiments (Fig 2.17 and 2.24), we have shown that *Ls-Otx* is specifically up-regulated in the earliest regenerating stages of *L.sanguineus* worms, during both, anterior and posterior regeneration. Components of the CNS, especially the nerve cords, are thought to play important roles during regeneration of nemerteans (Coe, 1932). Interestingly, *Ls-Otx* is predominantly expressed at the level of the CNS both, in adult (Fig. 2.11) and in developing nemerteans (Fig. 2.9). In addition, in regeneration experiments, we have detected *Ls-Otx* expression at the level of the amputated nerve cord (Fig. 2.19, B1). Due to this distinct expression pattern, we have proposed that *Otx* could be generally involved in the regeneration process of nemerteans.

To further test this hypothesis, we decided to investigate *Otx* expression pattern in *L.lacteus*, a very close relative of *L.sanguineus*. In fact, *L.sanguineus* and *L.lacteus* are more closely related to each other than *L.sanguineus* and *L.viridis* are. Both nemerteans, *L.sanguineus* and *L.lacteus,* are almost identical in morphological attributes and share similar environmental niches. However, these two species respond very differently to amputation (Fig. 2.25). Almost any isolated body fragment of *L.sanguineus* (except the body region 1) can undergo both, anterior and posterior regeneration. In contrast, all body fragments of *L.lacteus*, which have been isolated behind the brain, are incapable of anterior regeneration. In fact, multiple examples of such intraphyletic variability of regeneration capacities have already been described (Needham, 1952). This raises the question of why some animals can regenerate missing body parts, while others, even close relatives, cannot.

We made the assumption that an *Otx* homologue from *L.lacteus*, if present in its genome, would probably have a homeobox sequence close to the one of *L.sanguineus*. We have indeed already shown that *Otx* homeobox sequences from close *Lineus* relatives present an extremely high degree of identity (Fig. 2.7). Therefore, we examined the *Ll-Otx* expression pattern by using a Dig-labeled anti-sense RNA probe of only the *Ls-Otx* homeobox sequence.

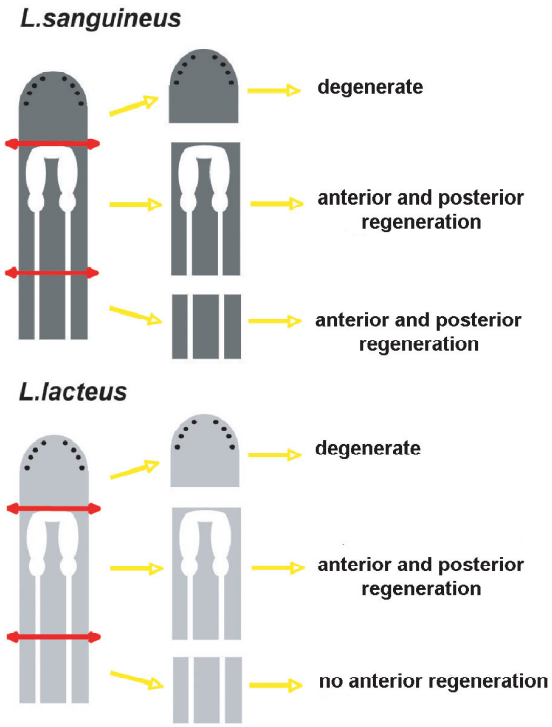


Fig. 2.25 The different regeneration properties of *L.sanguineus* and *L.lacteus*
When isolated, the body region 1 of both, *L.sanguineus* and *L.lacteus*, degenerates. Following amputation, all other body regions of *L.sanguineus* can undergo both, anterior and posterior regeneration. Posterior body fragments of *L.lacteus*, which result from an amputation above the brain, have also the capacity to undergo anterior and posterior regeneration. In contrast, if the amputation has been performed behind the brain, the resulting posterior body fragment of *L.lacteus* cannot undergo an anterior regeneration.

Similarly to the *Ls-Otx* expression in adult *L.sanguineus*, we detected expression of *Otx* in the adult brain of *L.lacteus* (not shown). Finding a similar expression pattern in *L.lacteus* than in *L.sanguineus* confirmed our assumption that we could use a *Ls-Otx* probe to investigate *Ll-Otx* expression.

In order to investigate *Ll-Otx* expression in amputated *L.lacteus*, we cut the body of several *L.lacteus* in the middle of region 4. In contrast to *L.sanguineus*, *L.lacteus* worms present a relatively big body region 4, which corresponds to the post-cerebral/ pre-esophageal region. Two body fragments result from such experiment: an "upper part", composed of body regions 1 to 4, which undergoes posterior regeneration and a "bottom part", composed of body regions 4 to 10, which cannot undergo anterior regeneration. We wanted to know whether *Otx* is expressed in this "bottom part", which is incapable of anterior regeneration. *Ll-Otx* is actually expressed at the amputation plane in such worm amputated from their body regions 1 to 4 (Fig. 2.26).

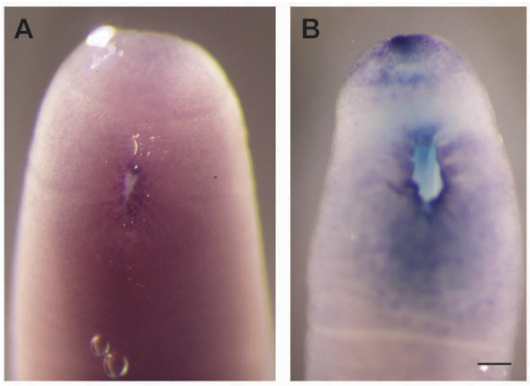


Fig. 2.26 *Ll-Otx* detection in amputated *L.lacteus*
ISH were done using *Ls-Otx* sense RNA probe in A and *Ls-Otx* anti-sense RNA probe in B on *L.lacteus* worms from which their body regions 1 to 4 have been amputated, five days before the fixation step with paraformaldehyde. Ventral view (notice the presence of the mouth), anterior orientated to the top. Notice the presence of background staining around the mouth.
Scale bar: 1,5 mm

We decided to explore the *Ll-Otx* expression in more detail at different time points after the amputation (Fig. 2.27), and to subsequently compare it to the *Ls-Otx* expression in similarly amputated *L.sanguineus* (Fig. 2.27, also see Fig. 2.19). In contrast to *L.lacteus*, *L.sanguineus* worms, which are missing their body regions 1 to 4, undergo anterior regeneration in order to restore their body pattern.

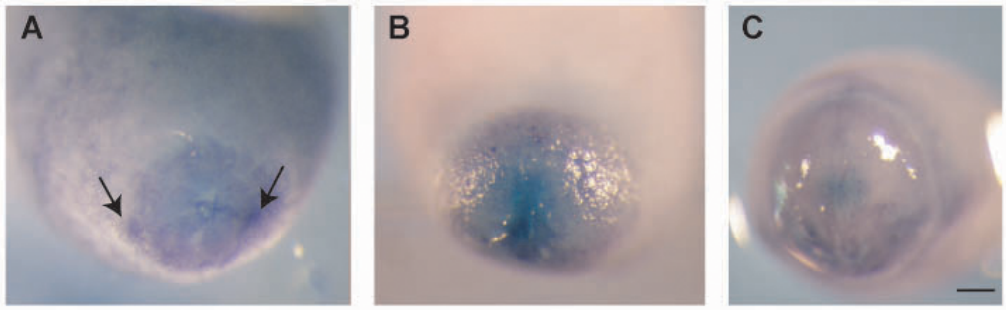


Fig. 2.27 *Ll-Otx* expression pattern at the plane of amputation
ISH were done using a *Ls-Otx* anti-sense RNA probe on *L.lacteus* worms, amputated from which their body regions 1 to 5 have been amputated. They are viewed from their anterior ends, which correspond to the plane of amputation.
Scale bar: 1 mm.
A: 3 dpa / B: 6 dpa / C: 10 dpa
Arrowheads in A indicate possible *Ll-Otx* expression in the nerve cord, at the level of their amputation plane.

Three days after the amputation, we observe some weak expression of *Ll-Otx* at the plane of amputation. In addition, we can distinguish two patches of cells showing a stronger *Ll-Otx* expression (Fig. 2.27, A). These patches of cells might correspond to the tip of the *L.lacteus* nerve cords, which have been sectioned by the amputation. Similar patches of cells, strongly expressing *Otx*, are also observed in corresponding *L.sanguineus* regenerating stages (Fig. 2.28, A and B). In these early regenerating stages, *Ls-Otx* is also expressed in an annular intermediate region at the amputation plane of *L.sanguineus* (Fig. 2.28, A and stronger in B) and later, at the level of large epithelial invaginations (Fig. 2.28, C). However, in *L.lacteus*, with the exception of the two patches of cells that might correspond to the tip of the sectioned lateral nerve cords, *Ll-Otx* expression pattern is rather diffuse, at the level of the amputation plane and does not present any corresponding characteristic with the *Ls-Otx* expression pattern, at similar *L.sanguineus* regenerating stages. Six days after the amputation, some *Ll-Otx* expression is still detectable at the plane of the amputation, in a median region (Fig. 2.27, B) and appears to become weaker, ten days after the amputation (Fig. 2.27, C).

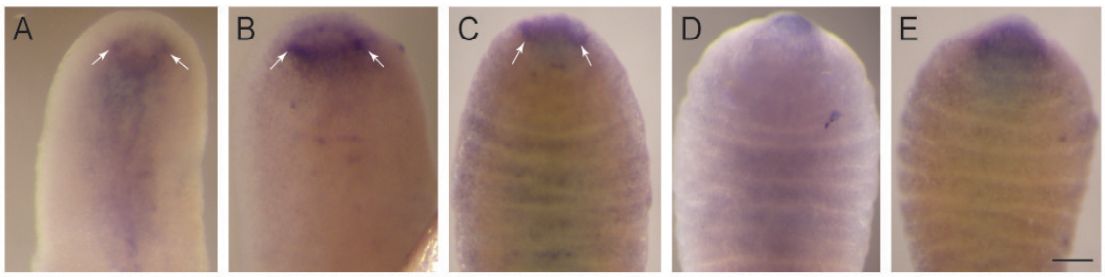


Fig. 2.28 *Otx* expression pattern in *L.sanguineus* during anterior regeneration
ISH were done using a *Ls-Otx* anti-sense RNA probe. Anterior regeneration from *L.sanguineus* missing body regions 1 to 5. Ventral view, anterior to the top. Scale bar: 1 mm.
A: 3 dpa / B: 4 dpa / C: 6 dpa / D: 8 dpa / E: 10 dpa.
White arrows indicate *Ls-Otx* expression in the nerve cords at the level of the amputation plane in A, B and C. Note the large epithelial invaginations at the level of the amputation plane in C.

4. DISCUSSION

4.1 Isolation and characterization of *Ls-Otx* and *Ls-Cdx*

4.1.1 The *Ls-Otx* gene

In order to isolate a *L.sanguineus Otx* homologue, we have conducted PCR with degenerated primers corresponding to different parts of the homeobox. We subsequently extended the obtained fragment by RACE PCR. This process led to the identification of a full-length *Ls-Otx* gene. We sequenced many clones containing putative Otx-like homeobox fragments, obtained by degenerated PCR. In every case, the sequence was identical. However, we cannot exclude the possibility that the *L.sanguineus* genome contains another *Otx* gene. Indeed, apparently independent *Otx* gene duplications have occurred several times during evolution in several lineages: in flatworms (Umesono *et al.*, 1999), in beetles (Li *et al.*, 1996), in lampreys (Ueki *et al.*, 1998) and in zebrafish (Mori *et al.*, 1994). The predicted Ls-Otx protein sequence is highly conserved. It contains a HD, which exhibits a lysine at the position 50 (K50), a characteristic that is shared with all members of the Otx family (Galliot *et al.*, 1999). This K50 distinguishes the Otx family from the Pax family (S50) and most other prd-related homeodomains (Q50). This lysine is critical for the DNA-binding specificity of the Otx members (Hanes and Brent, 1989), and confers specificity for the sequence TAATCC/T. The deduced Ls-Otx protein also contains another highly conserved motif, called WSP motif, in reference to the central three conserved aa: tryptophane (W)- serine (S)- proline (P) (Muller *et al.*, 1999). The WSP motif was defined as a motif of seven residues: SIWSPAS from the human CRX (Freund *et al.*, 1997). Six of these residues are conserved in the WSP motif of Ls-Otx (Fig. 2.6). Several Otx members also exhibit a hydrophobic C-terminal tail motif (Muller *et al.*, 1999), which has been duplicated in vertebrate sequences (Williams and Holland, 1998). This motif, which has been less conserved, is missing in Ls-Otx. When comparing the presence of those motifs and their sequences among various Otx members, a great heterogeneity appears: while humans and sea urchins Otx homologues possess both, a WSP motif and a tail motif, the beetle Otd2 possesses only the WSP motif and the leech Otx, only a degenerated tail motif. The *Drosophila* Otd sequence appears very derived, when compared to other Otx members: it lacks

both motifs and, in addition, is extensively longer (Muller *et al.*, 1999). It is known that, in addition to the HD, other functional motifs, present in homeodomain-containing proteins, play important roles. This has been highlighted by the finding that the hexapeptide motif displays a crucial role for formation of the HoxB1-Pbx1 heterodimer (Piper *et al.*, 1999). Furthermore, the observed functional differences of Otx1 and Otx2 in mice and in transgenic *Drosophila* have been proposed to rely on other domains, than the HD (Acampora *et al.*, 1995; Acampora *et al.*, 1996; Leuzinger *et al.*, 1998). It is known that the *Drosophila* Otd can only partially rescue the *Otx1*-deficient mouse phenotype (Acampora *et al.*, 1998). This could be due to the fact that, in contrast to the mouse Otx homologues, Otd is lacking both the WSP and tail motifs. The WSP motif and tail motif are functionally not well characterized. Further investigations will reveal whether their absence or presence might be at the basis of functional modulations of the different Otx homologues among bilaterian animals.

4.1.2 The *Ls-Cdx* gene

By using specific primers designed from the *Cdx*-type homeobox fragment already identified in *L.sanguineus* genome, we have isolated the full-length sequence of an *Ls-Cdx* gene. It contains a HD, which exhibits a Q50, as most prd-related homeodomains do. Immediately upstream of the HD, a five amino acid stretch, GKTRT, is present. Immediately downstream of it, there is a short stretch of KKK. These residues flanking the HD are conserved in most of the Cdx-related proteins (Gamer and Wright, 1993). In addition to the HD, Ls-Cdx contains a hexapeptide motif (PYDWMR), which is very similar to the one found in other Cdx members (Gamer and Wright, 1993). The vertebrate Cdx homologues possess two other conserved domains, called A and B, which are located in the amino-terminal part of the protein. Even though these domains are functionally not well characterized, it has been proposed that the domain A encodes a cytoplasmic export signal (Trinh *et al.*, 1999) and, adjacent to the B domain, a potent transactivation domain has been identified (Rings *et al.*, 2001). However these domains have been reported only from vertebrates Cdx sequences, so far. Accordingly, the domains A and B are lacking in Ls-Cdx.

4.2 The most anterior and posterior *Lineus* body regions are specificied by *Otx* and *Cdx* homologues, respectively

Otx and *Cdx*, two homeobox-containing genes, have been suggested to play important evolutionarily conserved roles in the early patterning of bilaterian embryos. Based on comparative studies in vertebrate and ecdysozoan embryos, it has been suggested that mutual interactions between *Otx* and *Cdx* homologues are part of a conserved mechanism used to establish early the A-P axis of the embryo (Isaacs *et al.*, 1999; Lynch *et al.*, 2006). Due to technical difficulties, the investigation of the expression patterns of *Otx* and *Cdx* homologues in the early developing stages of *L.viridis* was not easy. So far, the youngest developing stages from which we could obtain reproducible expression patterns are the 10 days-old larvae. As an A-P axis is already well-defined in these developing larval stages, we cannot hypothesize any involvement of *Otx* and *Cdx* in the early A-P axis establishment in *Lineus* from our gene expression data. Hence, it still remains unclear whether an interaction between *Otx* and *Cdx* homologues are part of the early A-P patterning system of a lophotrochozoan member.

But, in addition to their involvement in the early A-P axis establishment of Bilateria, *Otx* and *Cdx* homologues are also known to be key players in the specification of the most anterior and posterior body structures of bilaterian animals, respectively. These latter *Otx* and *Cdx* functions seem to be conserved in *Lineus*: while *Otx* is broadly expressed in almost all the developing head structures, including the brain, of *L.viridis* larvae, the expression of *Cdx* is specifically restricted to the most posterior part of the *L.viridis* larvae.

The fact that *Otx* is strongly expressed in the developing brain of a 10 days-old *L.viridis* larva is in good agreement with its assumed conserved role in neurogenesis and brain patterning among bilaterian animals. As it is also expressed in the adult brain, *Otx* seems to be involved in the maintenance of the identity of some brain territories in *L.sanguineus*. Finding such evolutionarily conserved *Otx* expression domains in the nemertean brain is in agreement with the idea of a monophyletic origin of the brain (Lichtneckert and Reichert, 2005). Furthermore, it

has been proposed that the bilaterian CNS is largely divided into three regions, each defined by specific gene expression domains (Reichert, 2005): while the anterior brain region is expressing *Otx,* the *Hox* genes are collinearly expressed in more posterior regions of the CNS. In addition, a domain of *Pax-2/5/8* expression is located between these two regions, the "non-*Hox*" one and the "*Hox"* one. As, in addition to the already available *LsHox* gene sequences, we have identified a *Ls-Otx* homologue and a *LsPax-2/5/8* homologue (See chapter III) during this PhD work, we are planning to test whether the tripartite brain organization is also conserved in a lophotrochozoan animal, by analyzing the *Ls-Otx, Ls-Pax2/5/8* and *LsHox* expression domains along the CNS of *L.sanguineus*. From our analysis, there is evidence that the *Ls-Otx* expression in the CNS is restricted to the anterior part of the brain. Although further investigations are needed, this is already in good agreement with the tripartite brain organization hypothesis.

The expression pattern of *Otx* is more restricted in the adult *Lineus* brain than in the developing one: during development, *Otx* is broadly expressed in both ventral and dorsal cerebral ganglia and their associated commissures whereas, in adult brains, it remains expressed only at the anterior periphery of the cerebral organs, on their external sides. Other genes are likely to specify the identity of other adult brain territories. Another highly conserved homeobox-containing gene, *Orthopedia* (*Otp*), is a good candidate to achieve this function in *Lineus* brain. This gene exemplifies in its name the fact that it encodes a homeodomain that shares similarities with the one encoded by both *Otd* and *Antp*. Despite differences in the expression patterns of *Otp* homologues in the mouse and in the fly, expression along the CNS of both animals has been observed**:** it is expressed in a metameric pattern along the brain and the ventral nerve cord of *Drosophila* (Simeone *et al.*, 1994) and, in mouse, it is expressed in some regions of the diencephalon, hindbrain and spinal cord (Simeone *et al.*, 1994). The mouse *Otp* displays one of its most important roles in the development of the hypothalamus (Acampora *et al.*, 1999; Wang and Lufkin, 2000). The expression patterns of lophotrochozoan *Otp* homologues suggest that *Otp* could also be involved in the CNS development of Lophotrochozoa: *Otp* expression has been associated with the development of the nervous system of the mollusk *Patella vulgata* (Nederbragt *et al.*, 2002) and with the specification of distinct regions of the adult brain of planaria *Dugesia tigrina* (Umesono *et al.*, 1999). In bilaterian animals, the expression of *Otx*

homologues always correlates with the development of the anterior nervous system, but, in fact, not with the development of the most anterior brain regions (Simeone *et al.*, 1992; Hirth *et al.*, 1995; Bruce and Shankland, 1998; Wada and Saiga, 1999). A comparable situation is observed in lophotrochozoan animals. Although the two planarian *Otx* homologues are expressed in almost all the brain of *Dugesia tigrina*, they are absent from the most anterior structures of the adult brain, the "so-called" brain branches. These most anterior brain structures actually express an *Otp* homologue (Umesono *et al.*, 1999). Similarly, *Otp* is found expressed in the most anterior nervous structures of the *Patella vulgata* larva, where *Otx* is not expressed. Hence, it seems that in these two lophotrochozoan species, *Otx* and *Otp* are expressed in complementary and non-overlapping regions of the anterior brain. From expression data in the youngest *L.viridis* larvae and in adult *L.sanguineus*, the medial anterior part of the cerebral ganglia seems to lack *Otx* expression. There must be another gene responsible for the development and maintenance of the identity of this restricted *Lineus* brain area. Based on lophotrochozoan *Otp* expression patterns, it would be interesting to search for a *L.sanguineus Otp* homologue. We could test, by double ISH experiment, whether this apparent mutual exclusion of the *Otx* and *Otp* expression is also observed in the development and maintenance of the *Lineus* brain. In *Lineus*, *Otx* seems to be involved in the specification of several anterior sensory and secretory organs, such as the cephalic gland, the frontal organ and the cerebral organs, which are assumed to display endocrine functions among others. *Otx* could be associated with the innervation process of these organs as it is strongly expressed in the nerve fibers that emerge from the brain to innervate the anterior sensory and secretory organs. Interestingly, these nerve fibers emerge from the territories of the adult *L.sanguineus* brain that retain some *Otx* expression. Similarly, *Otx* carries out conserved roles in the development of several sensory organs, such as the vertebrate inner ear, some sensory pigment cells in ascidians and also in secretory organs with endocrine function, such as the vertebrate pineal gland (Morsli *et al.*, 1999; Sanchez-Calderon *et al.*, 2002; Nishida *et al.*, 2003; Wada *et al.*, 2004). Thus, from our data, the important function of *Otx* in the development and innervation of deuterostomian and ecdysozoan sensory and secretory organs could be conserved in lophotrochozoan animals. From planarians to flies and vertebrates, *Otx* homologues are known to display an apparent evolutionary conserved role in the development of visual structures (Vandendries *et al.*, 1996; Umesono *et al.*, 1999; Nishida *et al.*, 2003). This

role does not seem to be conserved in *Lineus* since we never observed any expression pattern that could suggest an involvement of *Otx* in the development or in the maintenance of the adult *Lineus* eyes. In fact, a correlation of a precise localization of *Otx* expression with the location of the developing *Lineus* eyes is not observed, neither during embryonic development nor during regeneration. However, as *Otx* is broadly expressed in the anterior part of the *L.viridis* larvae, in the brain and in the nerve fibers emerging from it, we can still hypothesize for an early role of *Otx* in the "preparation" of the *Lineus* eye field. The *Lineus* eyes actually develop from a region that was, or it still (as it is difficult to know exactly when the eyes are forming) expressing *Otx*. This is a general feature of the bilaterian eyes, which develop in a body region that has been preliminary defined by *Otx.* In agreement with this notion, mice deficient in *Otx2* lack eyes because of the absence of forebrain, from which the eye field originate (Acampora *et al.*, 1995). It is also interesting to notice that in the fish, ectopic eye tissues can only be generated in the head region, which has been defined by *Otx* expression (Chuang and Raymond, 2002). We proposed that *Otx* has no direct role in *Lineus* eye specification, but rather a role in the specification of some head territories from which eyes will develop latter, possibly through coordinated function of several transcription factors, such as *Pax-6, Six1/2* and *Dac* (See chapter III).

As it is always expressed in the most posterior part of the developing bilaterian embryo, the *Cdx* gene has been proposed to be an ancestral master organizer of the posterior part of the bilaterian embryos and of the patterning along the A-P bilaterian body axis (Macdonald and Struhl, 1986; Gamer and Wright, 1993; Marom *et al.*, 1997; Katsuyama *et al.*, 1999; Moreno and Morata, 1999; Le Gouar *et al.*, 2003). But its presumed ancestral function in the early establishment of the A-P body axis has been challenged by data from *Drosophila*, where *Cad* acts only as a pair-rule gene regulator. However very recently, it has been shown that, in *Nasonia, Cad* plays a greater role in patterning the embryo than it does in *Drosophila* and that its function extends more anteriorly. In addition, whereas *Cad* mostly regulates the expression of pair-rule genes in *Drosophila*, in *Nasonia,* it activates the expression of gap genes, which then activate the expression of pair-rule genes (Olesnicky *et al.*, 2006). This places *Cad* at the top of the segmentation network in *Nasonia* and reinforces the idea that, in dipteran insects, like *Drosophila,* the role of *Cad* has been

reduced during evolution. The fact that *Cdx* is specifically expressed at the posterior end of the developing *L.viridis* larvae is in good agreement with the idea that *Cdx* homologues display an evolutionarily conserved role in the patterning of the posterior part of the bilaterian embryos. From its expression pattern in the developing *L.viridis*, we have suggested a possible involvement of *Cdx* in specification and development of the most caudal part of the *Lineus* gut. Our hypothesis is in line with the fact that *Cdx* homologues are known to play crucial roles in the patterning of the gut among bilaterian animals. For instance, in addition to its early role in the patterning of the posterior segments, *cad* is necessary for the proper development and maintenance of the hindgut primordium and proper specification of anal structures in *Drosophila* (Lengyel and Iwaki, 2002). Interestingly, *Cdx2* +/- mice exhibit multiple polyps that contain forestomach epithelium in their midgut. This suggests that a homeotic transformation that involves the gut endoderm has occurred in such mutant mice and has lead to the ectopic formation of anterior structures (= forestomach epithelium) at most posterior places (= midgut) (Chawengsaksophak *et al.*, 1997). In contrast, transgenic mice, which ectopically express *Cdx2* in their stomach, exhibit specific intestinal cell-types in their stomach (Mutoh *et al.*, 2002; Silberg *et al.*, 2002).

The *Cdx* expression pattern in developing *L.viridis* could also argue for an involvement in axial elongation of the body by posterior growth. This *Cdx* function is thought to be highly conserved among bilaterian animals: in addition to defects in the A-P patterning mediated through the *Hox* genes, *Cdx* mutant mice and morphant zebrafish show a truncated tail (van den Akker *et al.*, 2002; Chawengsaksophak *et al.*, 2004; Shimizu *et al.*, 2005). Similarly, RNAi on *Cdx* in the short germ band *Tribolium*, in the intermediate germ band *Gryllus* and in the crustacean *Artemia* have resulted in severe phenotypes of posterior body truncation (Copf *et al.*, 2004; Shinmyo *et al.*, 2005). This has lead to the idea that the A-P axis formation of the Urbilateria relied on an eventual posterior growth zone expressing *Cdx* (de Rosa *et al.*, 2005). Hence, the posterior part of the *L.viridis* larvae that expresses *Cdx* may correspond to a posterior growth zone. We have recently adapted a BrdU staining protocol to *Lineus*: As BrdU is a modified nucleotide that is incorporated into DNA during its replication, the detection of BrdU permits to reveal cells that are entering S-phase. Hence, such BrdU incorporation experiment in *L.viridis* larvae could allows us to visualize the part of the body where cell proliferation occurs. This would be interesting, using double detection of BrdU incorporation and *Cdx* expression, to test

whether the *Cdx* expressing posterior part of the *L.viridis* larvae also corresponds to a region where cell proliferation takes place. If *Cdx* is truly playing a role in *Lineus* axial elongation mediated by posterior growth, we expect to find some *Cdx* expression in adult *L.sanguineus* as well, as these animals continuously grown throughout their adult life. Unexpectedly, based on ISH results, *Cdx* is not expressed in adult *L.sanguineus*. However, by real time PCR analysis, we have detected some *Cdx* expression in the adult gut of *L.sanguineus*. This suggests that *Cdx* remains expressed in the differentiated gut but at too low a level for being detected by an ISH method. It has been suggested that mammalian *Cdx* homologues contribute to the cell renewal mechanism and to the control of stem cell differentiation in mature mammalian gut as they remained expressed in the differentiated gut (Beck, 2004). Likewise, we can hypothesize that *Cdx* plays a role in the regulation of intestinal cell proliferation, apoptosis, differentiation and dedifferentiation, by remaining expressed at low level in the differentiated intestine of *L.sanguineus*. These important mechanisms are probably tightly controlled in *L.sanguineus*, which has the ability to shrink in body mass by selectively reducing the "non-essential" body parts for its survival, such as the gut, under starvation conditions.

4.3 *Ls-Cdx* is specifically involved in the posterior regeneration

Ls-Cdx is strongly up-regulated within 24 hours in an isolated intestinal fragment, but only at the posterior end of the regenerating fragment. Hence, *Ls-Cdx* seems to play a role during the posterior regeneration of this isolated intestinal part: it could be involved in wound healing of the injured intestine and/ or regeneration of the missing most caudal part of it. In respect with this view, it is noteworthy that a strong up-regulation of *Cdx* homologues is observed in the damaged mammalian intestinal epithelium that undergoes regeneration (Subramanian *et al.*, 1998). From experiments that involve regeneration from isolated intestinal fragments, it is difficult to hypothesize any more general involvement of *Ls-Cdx* in the posterior regeneration of *L.sanguineus*: as we have indeed assumed a role of *Ls-Cdx* in specification of the caudal part of the gut, it is likely that *Ls-Cdx* would have a similar role in its regeneration. Hence, finding *Ls-Cdx* expressed at the posterior end of an isolated intestinal fragment that undergoes regeneration is not instructive for a putative more general involvement of *Ls-Cdx* in *L.sanguineus* posterior regeneration.

More evidences are indeed coming from worms that have been amputated of all the posterior structures that are located behind the cerebral organs: in such worms, *Ls-Cdx* is expressed at the level of the posterior blastema, from which an esophageal region will regenerate first. This means that *Ls-Cdx* is not only involved in the regeneration and patterning of the intestine, but probably plays a more general role in the posterior regeneration. As it is never expressed at the level of an anterior blastema, we conclude that *Ls-Cdx* is not involved in the anterior regeneration. Hence, the specific up-regulation of *Ls-Cdx* during posterior regeneration argues for its involvement in the early patterning events that take place during regeneration.

4.4 *Otx* displays dual roles in *L.sanguineus* regeneration

By real-time PCR, we have shown that, in an isolated body fragment from intestinal region, which is undergoing regeneration on both of its extremities, the activation of *Ls-Otx* occurs in two phases: within the first three days, there is an up-regulation of its expression, which is followed by a weak decrease in its expression. And then, after the first week post fragment isolation, a constant *Ls-Otx* up-regulation takes place, which is stronger than the first one observed during the first three days. Interestingly, the ISH experiments have revealed that *Ls-Otx* is expressed at both, anterior and posterior regenerating extremities during the first week of regeneration. In contrast, during the following days, it remains specifically expressed only at the level of the anterior blastema. We have proposed that the two *Ls-Otx* waves of expression during the regeneration of an isolated intestinal fragment may correspond to two different roles of *Ls-Otx* in *L.sanguineus* regeneration. Additionally, we have shown that in a worm, which undergoes only posterior regeneration as a result of an amputation behind its mouth, *Ls-Otx* is also quickly up-regulated. However, this *Ls-Otx* expression, in contrast to the *Ls-Cdx* one, is transient and disappears within a few days. This suggests that *Ls-Otx* is not only involved in anterior regeneration of *L.sanguineus* but, also in the first steps of posterior regeneration.

4.4.1 General involvement of *Ls-Otx* in the onset of regeneration

It is tempting to hypothesize that *Ls-Otx* plays first a role in the onset of regeneration at both extremities, the anterior and the posterior one. Then, secondly, it plays a role in the specification of anterior structures only. As we know the importance of the CNS in nemertean regeneration (see General introduction), it is noteworthy that the early *Ls-Otx* expression is observed at the tip of the sectioned lateral nerve cords, at both ends. A general involvement of *Ls-Otx* for the regeneration of the CNS itself is unlikely as its posterior expression at the level of the sectioned nerve cords disappears within few days, even though posterior regeneration of the CNS will take place. It has been suggested that, in response to body injury, the *Lineus* CNS could either emit signals responsible for regeneration, or integrate such signals or both. Based on its early expression at the level of sectioned nerve cords, *Ls-Otx* seems to be a good candidate for being a member of such a signaling network responsible for both anterior and posterior *L.sanguineus* regeneration. Similarly to our observations, *DtOtx* is activated in both, anterior and posterior blastemata during regeneration of the planarian *Dugesia tigrina* (Stornaiuolo *et al.*, 1998). However some differences exist between the expression of *Ls-Otx* and *DtOtx* during posterior regeneration of amputated *L.sanguineus* and *D.tigrina*, respectively. While *Ls-Otx* is expressed in the posterior blastema during the first days of regeneration, *DtOtx* remains expressed throughout the process of posterior regeneration (Stornaiuolo *et al.*, 1998). In addition, *DtOtx* is expressed with a clear antero-posterior asymmetric pattern during regeneration: more transcripts are detected in anterior blastemata than in posterior ones (Stornaiuolo *et al.*, 1998). In contrast, *Ls-Otx* seems to be expressed at similar levels in both, anterior and posterior regenerating ends. The expression of *Otx* during regeneration of another planarian, *Dugesia japonica*, has been studied: astonishingly, *Otx* expression has been detected only during anterior regeneration, and has not been found during posterior regeneration (Umesono *et al.*, 1999). These *Otx* expression differences in posterior regeneration of these two planarian species suggest that their *Otx* homologues display different functions. Hence, *Otx* functions could have diverged in the planarian lineage. However, this hypothesis needs further analysis.

Besides its expression at the tip of the sectioned lateral nerve cords, *Ls-Otx* is also expressed at the level of the large invaginations, which are characteristic features of early regenerative stages. It has been proposed that some selected epithelial cells translocate from the surface, through these large epithelial invaginations to the inner, actively reorganized, blastema region. Hence, in addition to a presumed role in the onset of regeneration, this specific expression argues for a possible role of *Ls-Otx* in cell mobility and cell rearrangement in early regenerative stages of *L.sanguineus*. A role of *Otx* in cell movements has also been proposed in cnidaria and in vertebrates (Boncinelli and Mallamaci, 1995; Bally-Cuif and Boncinelli, 1997; Smith *et al.*, 1999; Yanze *et al.*, 1999). Hence, *Otx* homologues might have an evolutionarily conserved function in cell movement and rearrangement.

4.4.2 Specific involvement of *Ls-Otx* in anterior regeneration

In addition to its probable role in the onset of regeneration, *Ls-Otx* is probably involved later in the specification and patterning of anterior structures, including the brain. Similarly, the expression of *Otx* homologues has been associated with the development of anterior structures in regenerating annelids and planaria (Stornaiuolo *et al.*, 1998; Umesono *et al.*, 1999; Bely and Wray, 2001). This second presumed *Ls-Otx* function occurs of course only during anterior regeneration. This explains the decrease and final absence of *Ls-Otx* expression in posterior blastema after several days of regeneration, when the onset of regeneration has been achieved. It seems that *Ls-Otx* expression is not required for the occurrence of growth, differentiation and patterning events during posterior regeneration but seems to be involved for their occurrence during anterior regeneration. Contrastingly, *Ls-Cdx* expression is not required for the occurrence of such events during anterior regeneration, but is probably involved in their occurrence during posterior regeneration.

4.5 Development and regeneration use similar molecular mechanisms but not always identical ones.

The expression of *Ls-Otx* during anterior regeneration suggests an involvement in the regeneration of the cerebral ganglia and cerebral organs. Based on expression studies in *L.viridis* larvae, we propose a similar involvement of *Otx* in the

development of these organs. This similar presumed role of *Otx* during development and regeneration of *Lineus* correlates with the idea that molecular mechanisms involved in the re-development of various tissues and organs during regeneration are a recapitulation of those occurring in development. In agreement with this idea, several studies have shown that some genes are expressed in the same way during development and regeneration (Loosli *et al.*, 1996; Stark *et al.*, 1998; Cadinouche *et al.*, 1999). However, regeneration and development should not be considered as similar processes as they differ in several respects: in fact, in contrast to development, regeneration is initiated by numerous cells that can arise from differentiated tissues. Furthermore, newly regenerated tissues and organs must be integrated to the already existing ones, in respect with the pre-established body axes. Interestingly, some dissimilarities have been reported between the timing and localization of the expression of same genes during development or regeneration (Akimenko *et al.*, 1995; Gardiner *et al.*, 1995; Bely and Wray, 2001). Similar differences in *Lineus* gene expressions are sometimes observed between development and regeneration: for example, despite its expression in developing frontal organs, *Ls-Otx* is not expressed in regenerating ones. These data point at the important notion that development and regeneration processes often use similar molecular mechanisms but not always identical ones.

4.6 The different regeneration capacities of *L.sanguineus* and *L.lacteus* are reflected in differences in the expression pattern of *Otx* after amputation

We have shown that an amputation in *L.sanguineus* leads to a specific activation of *Otx* expression at the level of the amputation plane. From its expression pattern, we have suggested that *Otx* could be part of a signaling network that would be responsible for the occurrence of regeneration in *L.sanguineus*. According to this model, *Otx* would not be activated at the level of the amputation plane of a body fragment incapable of regeneration. To test this hypothesis, we have used a close relative of *L.sanguineus*, *L.lacteus*, which presents different regeneration capacities (See Fig. 2.23): a *L.lacteus* worm, which has been amputated from its body regions 1 to 5, is not capable of anterior regeneration. We found that, event though no regeneration will occur, *Otx* is broadly expressed at the level of the amputation plane of such a *L.lacteus* worm, during the first week post amputation. Although this worm

is not able to undergo regeneration, we consider the possibility that *Otx* is activated during *L.lacteus* wound healing, a process that requires cell movement and rearrangement. From previous experiments, we have already suggested a role of *Otx* in cell movement and rearrangement. Interestingly, three days after the amputation, *Otx* is also expressed at the level of the sectioned lateral nerve of *L.lacteus*, like it is expressed in a similarly amputated *L.sanguineus*. Hence, although *L.lacteus*, which has been amputated from its body regions 1 to 5, cannot undergo anterior regeneration, *Otx* is expressed at the tip of its sectioned nerve cords. As we have proposed that *Otx* expression at the level of the *L.sanguineus* sectioned nerve cords could be, at least in part, responsible for regeneration to occur, it is difficult to interpret this specific expression pattern of *Otx* in these amputated *L.lacteus* that are incapable of regeneration.

A grafting experiment between *L.sanguineus* and *L.lacteus*, as carried out by J.Bierne and M.Tarpin, offers an attractive explanation (J.Bierne and M.Tarpin, unpublished). They have first isolated an antecerebral *L.sanguineus* end, which is the single *L.sanguineus* body region that, when isolated, cannot regenerate due the absence of CNS cells. Then, they have grafted it onto a *L.lacteus* mid-body fragment, which is incapable of undergoing anterior regeneration. Once the graft was healed, they re-amputated the previously grafted *L.sanguineus* antecerebral end. They left only a minimal amount of *L.sanguineus* tissue on the *L.lacteus* recipient. Amazingly, under these experimental conditions, a complete worm is regenerated (Fig. 2.29). The regenerated worm displays all the characteristics of a *L.sanguineus* worm, such as the pigmentation, anatomical features and regeneration ability. Hence, the few adult cells from the antecerebral end of *L.sanguineus* are capable of regeneration, under specific experimental conditions. J.Bierne and M.Tarpin proposed that, in this grafting experiment, the sectioned *L.lacteus* nerve cords have induced the activation of the few adult *L.sanguineus* cells into *L.sanguineus* “stem cells” (personal communication). In addition, they have suggested that the *L.lacteus* recipient has provided, via “nurse cells”, the energy necessary for all the mechanisms, such as the conversion of adult *L.sanguineus* cells into “stem cells”, the blastema formation, cell proliferation, apoptosis and differentiation, which have to occur in order complete regeneration takes place.

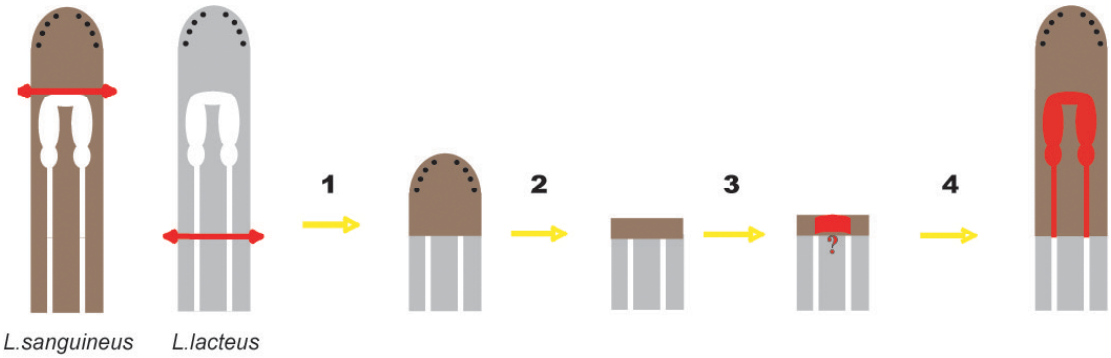


Fig. 2.29 Scheme of the grafting experiment that leads to regeneration of a complete worm form very few adult *L.sanguineus* cells originating from the antecerebral end

1: the antecerebral end of *L.sanguineus*, which is incapable of regeneration, is grafted on a *L.lacteus* mid-body fragment, which cannot regenerate a head.
2: the *L.sanguineus* antecerebral end is re-amputated. Only a minimal amount of *L.sanguineus* cells are left on the *L.lacteus* recipient.
3: several crucial mechanisms, such as wound healing, activation of cells from the dermis and epidermis into stem cells, neurogenesis and organogenesis of the various body components, take place…
4: … in order to achieve the regeneration of a complete worm.

Furthermore, in the reverse experiment (= the graft of a *L.lacteus* antecerebral end onto *L.sanguineus* mid-body fragment, and its subsequent removal), it has been shown that adult *L.lacteus* cells are not converted into *L.lacteus* stem cells by interaction with the sectioned *L.sanguineus* nerve cords. To summarize, these grafting experiments suggest that the sectioned *L.lacteus* nerve cords have the ability to induce the de-differentiation of adult *L.sanguineus* cells into stem cells.

In the light of these grafting experiments, we propose that, following their sectioning, the nerve cords of both, *L.sanguineus* and *L.lacteus*, initiate a signaling network that leads to regeneration. The fact that *Otx* is expressed at the tips of the sectioned *L.lacteus* nerve cords reinforces our idea that it is part of this signaling network. According to our model, the differences of the regeneration capacities between *L.sanguineus* and *L.lacteus* rely on the differences in the capacity of their differentiated cells to de-differentiate in response to such signals rather than in the capacity of their sectioned nerve cords to emit the signals. Interestingly, we have noticed that, during *L.sanguineus* regeneration, *Otx* is firstly expressed at the tips of the sectioned nerve cords and later, expressed in cells located in an annular region,

around the nerve cords. These cells translocate from the surface to the inside, actively reorganizing part of the blastema. It is tempting to hypothesize that these adult *L.sanguineus* cells that are/ or will be converted into stem cells express *Otx* in response to the signals emitted by the sectioned nerve cords. Remarkably, we never observed this secondary *Otx* expression in *L.lacteus*, which cannot regenerate anteriorly. This is in good agreement with the idea that adult *L.lacteus* cells cannot integrate the signals coming from their sectioned nerve cords.

5. References of chapter II

Acampora, D., Mazan, S., Lallemand, Y., Avantaggiato, V., Maury, M., Simeone, A. and Brulet, P. (1995). "Forebrain and midbrain regions are deleted in *Otx2-/-* mutants due to a defective anterior neuroectoderm specification during gastrulation." *Development* **121**(10): 3279-3290.

Acampora, D., Mazan, S., Avantaggiato, V., Barone, P., Tuorto, F., Lallemand, Y., Brulet, P. and Simeone, A. (1996). "Epilepsy and brain abnormalities in mice lacking the *Otx1* gene." *Nat Genet* **14**(2): 218-222.

Acampora, D., Avantaggiato, V., Tuorto, F. and Simeone, A. (1997). "Genetic control of brain morphogenesis through *Otx* gene dosage requirement." *Development* **124**(18): 3639-3650.

Acampora, D., Avantaggiato, V., Tuorto, F., Barone, P., Reichert, H., Finkelstein, R. and Simeone, A. (1998). "Murine *Otx1* and *Drosophila otd* genes share conserved genetic functions required in invertebrate and vertebrate brain development." *Development* **125**(9): 1691-1702.

Acampora, D., Postiglione, M. P., Avantaggiato, V., Di Bonito, M., Vaccarino, F. M., Michaud, J. and Simeone, A. (1999). "Progressive impairment of developing neuroendocrine cell lineages in the hypothalamus of mice lacking the *Orthopedia* gene." *Genes Dev* **13**(21): 2787-2800.

Acampora, D., Boyl, P. P., Signore, M., Martinez-Barbera, J. P., Ilengo, C., Puelles, E., Annino, A., Reichert, H., Corte, G. and Simeone, A. (2001). "OTD/OTX2 functional equivalence depends on 5' and 3' UTR-mediated control of *Otx2* mRNA for nucleo-cytoplasmic export and epiblast-restricted translation." *Development* **128**(23): 4801-4813.

Adachi, Y., Nagao, T., Saiga, H. and Furukubo-Tokunaga, K. (2001). "Cross-phylum regulatory potential of the ascidian *Otx* gene in brain development in *Drosophila melanogaster*." *Dev Genes Evol* **211**(6): 269-280.

Akimenko, M. A., Johnson, S. L., Westerfield, M. and Ekker, M. (1995). "Differential induction of four msx homeobox genes during fin development and regeneration in zebrafish." *Development* **121**(2): 347-357.

Altmann, C. R. and Brivanlou, A. H. (2001). "Neural patterning in the vertebrate embryo." *Int Rev Cytol* **203**: 447-482.

Ang, S. L., Jin, O., Rhinn, M., Daigle, N., Stevenson, L. and Rossant, J. (1996). "A targeted mouse *Otx2* mutation leads to severe defects in gastrulation and formation of axial mesoderm and to deletion of rostral brain." *Development* **122**(1): 243-252.

Arendt, D. and Nubler-Jung, K. (1996). "Common ground plans in early brain development in mice and flies." *Bioessays* **18**(3): 255-259.

Arendt, D. and Nubler-Jung, K. (1999). "Comparison of early nerve cord development in insects and vertebrates." *Development* **126**(11): 2309-2325.

Bally-Cuif, L. and Boncinelli, E. (1997). "Transcription factors and head formation in vertebrates." *Bioessays* **19**(2): 127-135.

Beck, F. (2004). "The role of *Cdx* genes in the mammalian gut." *Gut* **53**(10): 1394-1396.

Bely, A. E. and Wray, G. A. (2001). "Evolution of regeneration and fission in annelids: insights from *engrailed*- and *orthodenticle*-class gene expression." *Development* **128**(14): 2781-2791.

Blitz, I. L. and Cho, K. W. (1995). "Anterior neurectoderm is progressively induced during gastrulation: the role of the *Xenopus* homeobox gene *orthodenticle*." *Development* **121**(4): 993-1004.

Boncinelli, E. and Mallamaci, A. (1995). "Homeobox genes in vertebrate gastrulation." *Curr Opin Genet Dev* **5**(5): 619-627
.
Boncinelli, E. and Morgan, R. (2001). "Downstream of Otx2, or how to get a head." *Trends Genet* **17**(11): 633-636.

Brooke, N. M., Garcia-Fernandez, J. and Holland, P. W. (1998). "The ParaHox gene cluster is an evolutionary sister of the *Hox* gene cluster." *Nature* **392**(6679): 920-922.

Bruce, A. E. and Shankland, M. (1998). "Expression of the head gene *Lox22-Otx* in the leech *Helobdella* and the origin of the bilaterian body plan." *Dev Biol* **201**(1): 101-112.

Cadinouche, M. Z., Liversage, R. A., Muller, W. and Tsilfidis, C. (1999). "Molecular cloning of the *Notophthalmus viridescens radical fringe* cDNA and characterization of its expression during forelimb development and adult forelimb regeneration." *Dev Dyn* **214**(3): 259-268.

Castro, L. F., Rasmussen, S. L., Holland, P. W., Holland, N. D. and Holland, L. Z. (2006). "A *Gbx* homeobox gene in amphioxus: insights into ancestry of the ANTP class and evolution of the midbrain/hindbrain boundary." *Dev Biol* **295**(1): 40-51.

Charite, J., de Graaff, W., Consten, D., Reijnen, M. J., Korving, J. and Deschamps, J. (1998). "Transducing positional information to the *Hox* genes: critical interaction of *cdx* gene products with position-sensitive regulatory elements." *Development* **125**(22): 4349-4358.

Chawengsaksophak, K., James, R., Hammond, V. E., Kontgen, F. and Beck, F. (1997). "Homeosis and intestinal tumours in *Cdx2* mutant mice." *Nature* **386**(6620): 84-87.

Chawengsaksophak, K., de Graaff, W., Rossant, J., Deschamps, J. and Beck, F. (2004). "*Cdx2* is essential for axial elongation in mouse development." *Proc Natl Acad Sci U S A* **101**(20): 7641-7645.

Chuang, J. C. and Raymond, P. A. (2002). "Embryonic origin of the eyes in teleost fish." *Bioessays* **24**(6): 519-529.

Coe, W. R. (1932). "Regeneration in nemerteans. III Regeneration of *Lineus pictifrons*." *J. Exp. Zool* **61**: 29-43.

Cohen, S. M. and Jurgens, G. (1990). "Mediation of *Drosophila* head development by gap-like segmentation genes." *Nature* **346**(6283): 482-485.

Copf, T., Schroder, R. and Averof, M. (2004). "Ancestral role of *caudal* genes in axis elongation and segmentation." *Proc Natl Acad Sci U S A* **101**(51): 17711-17715.

Cornell, R. A. and Ohlen, T. V. (2000). "*Vnd/nkx*, *ind/gsh*, and *msh/msx*: conserved regulators of dorsoventral neural patterning?" *Curr Opin Neurobiol* **10**(1): 63-71.

Davis, G. K. and Patel, N. H. (2002). "Short, long, and beyond: molecular and embryological approaches to insect segmentation." *Annu Rev Entomol* **47**: 669-699.

De Robertis, E. M. and Sasai, Y. (1996). "A common plan for dorsoventral patterning in Bilateria." *Nature* **380**(6569): 37-40.

De Rosa, R., Prud'homme, B. and Balavoine, G. (2005). "*Caudal* and *even-skipped* in the annelid *Platynereis dumerilii* and the ancestry of posterior growth." *Evol Dev* **7**(6): 574-587.

Delsuc, F., Brinkmann, H., Chourrout, D. and Philippe, H. (2006). "Tunicates and not cephalochordates are the closest living relatives of vertebrates." *Nature* **439**(7079): 965-968.

Deschamps, J. and van Nes, J. (2005). "Developmental regulation of the *Hox* genes during axial morphogenesis in the mouse." *Development* **132**(13): 2931-2942.

Driever, W. and Nusslein-Volhard, C. (1988). "The bicoid protein determines position in the *Drosophila* embryo in a concentration-dependent manner." *Cell* **54**(1): 95-104.

Epstein, M., Pillemer, G., Yelin, R., Yisraeli, J. K. and Fainsod, A. (1997). "Patterning of the embryo along the anterior-posterior axis: the role of the *caudal* genes." *Development* **124**(19): 3805-3814.

Finkelstein, R., Smouse, D., Capaci, T. M., Spradling, A. C. and Perrimon, N. (1990a). "The *orthodenticle* gene encodes a novel homeodomain protein involved in the development of the *Drosophila* nervous system and ocellar visual structures." *Genes Dev* **4**(9): 1516-1527.

Finkelstein, R. and Perrimon, N. (1990b). "The *orthodenticle* gene is regulated by bicoid and torso and specifies *Drosophila* head development." *Nature* **346**(6283): 485-488.

Freund, C. L., Gregory-Evans, C. Y., Furukawa, T., Papaioannou, M., Looser, J., Ploder, L., Bellingham, J., Ng, D., Herbrick, J. A., Duncan, A., Scherer, S. W., Tsui, L. C., Loutradis-Anagnostou, A., Jacobson, S. G., Cepko, C. L., Bhattacharya, S. S. and McInnes, R. R. (1997). "Cone-rod dystrophy due to mutations in a novel photoreceptor-specific homeobox gene (*CRX*) essential for maintenance of the photoreceptor." *Cell* **91**(4): 543-553.

Frohnhöfer, H. G. and Nüsslein-Volhard, C. (1986). "Organization of anterior pattern int he *Drosophila* embryo by the maternal gene *bicoid*." *nature* **324**: 120-125.

Furukawa, T., Morrow, E. M. and Cepko, C. L. (1997). "*Crx*, a novel *otx*-like homeobox gene, shows photoreceptor-specific expression and regulates photoreceptor differentiation." *Cell* **91**(4): 531-541.

Galliot, B., de Vargas, C. and Miller, D. (1999). "Evolution of homeobox genes: Q50 Paired-like genes founded the Paired class." *Dev Genes Evol* **209**(3): 186-197.

Gallitano-Mendel, A. and Finkelstein, R. (1998). "Ectopic *orthodenticle* expression alters segment polarity gene expression but not head segment identity in the *Drosophila* embryo." *Dev Biol* **199**(1): 125-137.

Gamer, L. W. and Wright, C. V. (1993). "Murine *Cdx-4* bears striking similarities to the Drosophila caudal gene in its homeodomain sequence and early expression pattern." *Mech Dev* **43**(1): 71-81.

Gardiner, D. M., Blumberg, B., Komine, Y. and Bryant, S. V. (1995). "Regulation of *HoxA* expression in developing and regenerating axolotl limbs." *Development* **121**(6): 1731-1741.

Gibson, R., Ed. (1972). Nemerteans, *Hutchinson Publ. London*.

Hanes, S. D. and Brent, R. (1989). "DNA specificity of the bicoid activator protein is determined by homeodomain recognition helix residue 9." *Cell* **57**(7): 1275-1283.

Hirth, F., Therianos, S., Loop, T., Gehring, W. J., Reichert, H. and Furukubo-Tokunaga, K. (1995). "Developmental defects in brain segmentation caused by mutations of the homeobox genes *orthodenticle* and *empty spiracles* in *Drosophila*." *Neuron* **15**(4): 769-778.

Hirth, F., Hartmann, B. and Reichert, H. (1998). "Homeotic gene action in embryonic brain development of *Drosophila*." *Development* **125**(9): 1579-1589.

Hirth, F., Kammermeier, L., Frei, E., Walldorf, U., Noll, M. and Reichert, H. (2003). "An urbilaterian origin of the tripartite brain: developmental genetic insights from *Drosophila*." *Development* **130**(11): 2365-2373.

Hulskamp, M., Pfeifle, C. and Tautz, D. (1990). "A morphogenetic gradient of hunchback protein organizes the expression of the gap genes *Kruppel* and *knirps* in the early *Drosophila* embryo." *Nature* **346**(6284): 577-580.

Hunter, C. P. and Kenyon, C. (1996). "Spatial and temporal controls target pal-1 blastomere-specification activity to a single blastomere lineage in *C.elegans* embryos." *Cell* **87**(2): 217-226.

Isaacs, H. V., Andreazzoli, M. and Slack, J. M. (1999). "Anteroposterior patterning by mutual repression of orthodenticle and caudal-type transcription factors." *Evol Dev* **1**(3): 143-152.

Katsuyama, Y., Sato, Y., Wada, S. and Saiga, H. (1999). "Ascidian tail formation requires caudal function." *Dev Biol* **213**(2): 257-268.

Kmita-Cunisse, M., Loosli, F., Bierne, J. and Gehring, W. J. (1998). "Homeobox genes in the ribbonworm *Lineus sanguineus*: evolutionary implications." *Proc Natl Acad Sci U S A* **95**(6): 3030-3035.

Krumlauf, R. (1994). "*Hox* genes in vertebrate development." *Cell* **78**(2): 191-201.

Le Gouar, M., Lartillot, N., Adoutte, A. and Vervoort, M. (2003). "The expression of a *caudal* homologue in a mollusc, *Patella vulgata*." *Gene Expr Patterns* **3**(1): 35-37.

Lee, J. E. (1997). "Basic helix-loop-helix genes in neural development." *Curr Opin Neurobiol* **7**(1): 13-20.

Lengyel, J. A. and Iwaki, D. D. (2002). "It takes guts: the *Drosophila* hindgut as a model system for organogenesis." *Dev Biol* **243**(1): 1-19.

Leuzinger, S., Hirth, F., Gerlich, D., Acampora, D., Simeone, A., Gehring, W. J., Finkelstein, R., Furukubo-Tokunaga, K. and Reichert, H. (1998). "Equivalence of the fly *orthodenticle* gene and the human *OTX* genes in embryonic brain development of *Drosophila*." *Development* **125**(9): 1703-1710.

Lewis, J. (1996). "Neurogenic genes and vertebrate neurogenesis." *Curr Opin Neurobiol* **6**(1): 3-10.

Li, Y., Brown, S. J., Hausdorf, B., Tautz, D., Denell, R. E. and Finkelstein, R. (1996). "Two *orthodenticle*-related genes in the short-germ beetle *Tribolium castaneum*." *Dev. Genes Evol.* **206**: 35-45.

Li, J. Y. and Joyner, A. L. (2001). "*Otx2* and *Gbx2* are required for refinement and not induction of mid-hindbrain gene expression." *Development* **128**(24): 4979-4991.

Lichtneckert, R. and Reichert, H. (2005). "Insights into the urbilaterian brain: conserved genetic patterning mechanisms in insect and vertebrate brain development." *Heredity* **94**(5): 465-477.

Lohnes, D. (2003). "The Cdx1 homeodomain protein: an integrator of posterior signaling in the mouse." *Bioessays* **25**(10): 971-980.

Loosli, F., Kmita-Cunisse, M. and Gehring, W. J. (1996). "Isolation of a *Pax-6* homolog from the ribbonworm *Lineus sanguineus*." *Proc Natl Acad Sci U S A* **93**(7): 2658-2663.

Lynch, J. and Desplan, C. (2003). "'De-evolution' of *Drosophila* toward a more generic mode of axis patterning." *Int J Dev Biol* **47**(7-8): 497-503.

Lynch, J. A., Brent, A. E., Leaf, D. S., Pultz, M. A. and Desplan, C. (2006). "Localized maternal *orthodenticle* patterns anterior and posterior in the long germ wasp *Nasonia*." *Nature* **439**(7077): 728-732.

Macdonald, P. M. and Struhl, G. (1986). "A molecular gradient in early *Drosophila* embryos and its role in specifying the body pattern." *Nature* **324**(6097): 537-545.

Marin, F. and Puelles, L. (1994). "Patterning of the embryonic avian midbrain after experimental inversions: a polarizing activity from the isthmus." *Dev Biol* **163**(1): 19-37.

Marom, K., Shapira, E. and Fainsod, A. (1997). "The chicken *caudal* genes establish an anterior-posterior gradient by partially overlapping temporal and spatial patterns of expression." *Mech Dev* **64**(1-2): 41-52.

Matsuo, K., Yoshida, H. and Shimizu, T. (2005). "Differential expression of *caudal* and *dorsal* genes in the teloblast lineages of the oligochaete annelid *Tubifex tubifex*." *Dev Genes Evol* **215**(5): 238-247.

McGinnis, W. and Krumlauf, R. (1992). "Homeobox genes and axial patterning." *Cell* **68**(2): 283-302.

McGregor, A. P. (2005). "How to get ahead: the origin, evolution and function of bicoid." *Bioessays* **27**(9): 904-913.

McGregor, A. P. (2006). "Wasps, beetles and the beginning of the ends." *Bioessays* **28**(7): 683-686.

Mlodzik, M., Fjose, A. and Gehring, W. J. (1985). "Isolation of *caudal*, a *Drosophila* homeobox-containing gene with maternal expression, whose transcripts form a concentration gradient at the pre-blastoderm stage." *Embo J* **4**(11): 2961-2969.

Mlodzik, M. and Gehring, W. J. (1987). "Expression of the *caudal* gene in the germ line of *Drosophila*: formation of an RNA and protein gradient during early embryogenesis." *Cell* **48**(3): 465-478.

Mlodzik, M., Gibson, G. and Gehring, W. J. (1990). "Effects of ectopic expression of *caudal* during *Drosophila* development." *Development* **109**(2): 271-277.

Montalta-He, H., Leemans, R., Loop, T., Strahm, M., Certa, U., Primig, M., Acampora, D., Simeone, A. and Reichert, H. (2002). "Evolutionary conservation of otd/Otx2 transcription factor action: a genome-wide microarray analysis in *Drosophila.*" *Genome Biol* **3**(4): RESEARCH0015.

Moreno, E. and Morata, G. (1999). "*Caudal* is the *Hox* gene that specifies the most posterior *Drosophila* segment." *Nature* **400**(6747): 873-877.

Mori, H., Miyazaki, Y., Morita, T., Nitta, H. and Mishina, M. (1994). "Different spatio-temporal expressions of three otx homeoprotein transcripts during zebrafish embryogenesis." *Brain Res Mol Brain Res* **27**(2): 221-231.

Morsli, H., Tuorto, F., Choo, D., Postiglione, M. P., Simeone, A. and Wu, D. K. (1999). "Otx1 and Otx2 activities are required for the normal development of the mouse inner ear." *Development* **126**(11): 2335-2343.

Muller, P., Yanze, N., Schmid, V. and Spring, J. (1999). "The homeobox gene *Otx* of the jellyfish *Podocoryne carnea*: role of a head gene in striated muscle and evolution." *Dev Biol* **216**(2): 582-594.

Mutoh, H., Hakamata, Y., Sato, K., Eda, A., Yanaka, I., Honda, S., Osawa, H., Kaneko, Y. and Sugano, K. (2002). "Conversion of gastric mucosa to intestinal metaplasia in *Cdx2*-expressing transgenic mice." *Biochem Biophys Res Commun* **294**(2): 470-479.

Nagao, T., Leuzinger, S., Acampora, D., Simeone, A., Finkelstein, R., Reichert, H. and Furukubo-Tokunaga, K. (1998). "Developmental rescue of *Drosophila* cephalic defects by the human *Otx* genes." *Proc Natl Acad Sci U S A* **95**(7): 3737-3742.

Nederbragt, A. J., te Welscher, P., van den Driesche, S., van Loon, A. E. and Dictus, W. J. (2002). "Novel and conserved roles for *orthodenticle/ otx* and *orthopedia/ otp* orthologs in the gastropod mollusc *Patella vulgata*." *Dev Genes Evol* **212**(7): 330-337.

Needham, A. E. (1952). Regeneration and Wound-Healing, *New York: John Wiley and Sons, Inc.*

Nielsen (1995). *Animal evolution. Interrelationships of the Living Phyla.*, *Oxford University/ Oxford.*

Nishida, A., Furukawa, A., Koike, C., Tano, Y., Aizawa, S., Matsuo, I. and Furukawa, T. (2003). "*Otx2* homeobox gene controls retinal photoreceptor cell fate and pineal gland development." *Nat Neurosci* **6**(12): 1255-1263.

Northrop, J. L. and Kimelman, D. (1994). "Dorsal-ventral differences in *Xcad-3* expression in response to FGF-mediated induction in *Xenopus*." *Dev Biol* **161**(2): 490-503.

Olesnicky, E. C., Brent, A. E., Tonnes, L., Walker, M., Pultz, M. A., Leaf, D. and Desplan, C. (2006). "A *caudal* mRNA gradient controls posterior development in the wasp *Nasonia*." *Development* **133**(20): 3973-3982.

Pannese, M., Polo, C., Andreazzoli, M., Vignali, R., Kablar, B., Barsacchi, G. and Boncinelli, E. (1995). "The *Xenopus* homologue of *Otx2* is a maternal homeobox gene that demarcates and specifies anterior body regions." *Development* **121**(3): 707-720.

Pillemer, G., Epstein, M., Blumberg, B., Yisraeli, J. K., De Robertis, E. M., Steinbeisser, H. and Fainsod, A. (1998). "Nested expression and sequential downregulation of the *Xenopus caudal* genes along the anterior-posterior axis." *Mech Dev* **71**(1-2): 193-196.

Pinnell, J., Lindeman, P. S., Colavito, S., Lowe, C. and Savage, R. M. (2006). "The divergent roles of the segmentation gene *hunchback*." *Integr. Comp. Biol.* **46**: 519-532.

Piper, D. E., Batchelor, A. H., Chang, C. P., Cleary, M. L. and Wolberger, C. (1999). "Structure of a HoxB1-Pbx1 heterodimer bound to DNA: role of the hexapeptide and a fourth homeodomain helix in complex formation." *Cell* **96**(4): 587-597.

Pourquie, O. (2003). "Vertebrate somitogenesis: a novel paradigm for animal segmentation?" *Int J Dev Biol* **47**(7-8): 597-603.

Pownall, M. E., Tucker, A. S., Slack, J. M. and Isaacs, H. V. (1996). "*eFGF*, *Xcad3* and *Hox* genes form a molecular pathway that establishes the anteroposterior axis in *Xenopus*." *Development* **122**(12): 3881-3892.

Pultz, M. A., Pitt, J. N. and Alto, N. M. (1999). "Extensive zygotic control of the anteroposterior axis in the wasp *Nasonia vitripennis*." *Development* **126**(4): 701-710.

Raible, F. and Brand, M. (2004). "Divide et Impera--the midbrain-hindbrain boundary and its organizer." *Trends Neurosci* **27**(12): 727-734.

Reichert, H. (2002). "Conserved genetic mechanisms for embryonic brain patterning." *Int J Dev Biol* **46**(1): 81-87.

Reichert, H. (2005). "A tripartite organization of the urbilaterian brain: developmental genetic evidence from *Drosophila*." *Brain Res Bull* **66**(4-6): 491-494.

Rhinn, M. and Brand, M. (2001). "The midbrain--hindbrain boundary organizer." *Curr Opin Neurobiol* **11**(1): 34-42.

Rings, E. H., Boudreau, F., Taylor, J. K., Moffett, J., Suh, E. R. and Traber, P. G. (2001). "Phosphorylation of the serine 60 residue within the Cdx2 activation domain mediates its transactivation capacity." *Gastroenterology* **121**(6): 1437-1450.

Rivera-Pomar, R., Lu, X., Perrimon, N., Taubert, H. and Jackle, H. (1995). "Activation of posterior gap gene expression in the *Drosophila* blastoderm." *Nature* **376**(6537): 253-256.

Rivera-Pomar, R., Niessing, D., Schmidt-Ott, U., Gehring, W. J. and Jackle, H. (1996). "RNA binding and translational suppression by bicoid." *Nature* **379**(6567): 746-749.

Royet, J. and Finkelstein, R. (1995). "Pattern formation in *Drosophila* head development: the role of the *orthodenticle* homeobox gene." *Development* **121**(11): 3561-3572.

Sanchez-Calderon, H., Martin-Partido, G. and Hidalgo-Sanchez, M. (2002). "Differential expression of *Otx2*, *Gbx2*, *Pax2*, and *Fgf8* in the developing vestibular and auditory sensory organs." *Brain Res Bull* **57**(3-4): 321-323.

Schroder, R. (2003). "The genes *orthodenticle* and *hunchback* substitute for *bicoid* in the beetle *Tribolium*." *Nature* **422**(6932): 621-625.

Shimizu, T., Bae, Y. K., Muraoka, O. and Hibi, M. (2005). "Interaction of *Wnt* and *caudal*-related genes in zebrafish posterior body formation." *Dev Biol* **279**(1): 125-141.

Shinmyo, Y., Mito, T., Matsushita, T., Sarashina, I., Miyawaki, K., Ohuchi, H. and Noji, S. (2005). "*caudal* is required for gnathal and thoracic patterning and for posterior elongation in the intermediate-germband cricket *Gryllus bimaculatus*." *Mech Dev* **122**(2): 231-239.

Silberg, D. G., Sullivan, J., Kang, E., Swain, G. P., Moffett, J., Sund, N. J., Sackett, S. D. and Kaestner, K. H. (2002). "*Cdx2* ectopic expression induces gastric intestinal metaplasia in transgenic mice." *Gastroenterology* **122**(3): 689-696.

Simeone, A., Acampora, D., Gulisano, M., Stornaiuolo, A. and Boncinelli, E. (1992). "Nested expression domains of four homeobox genes in developing rostral brain." *Nature* **358**(6388): 687-690.

Simeone, A., Acampora, D., Mallamaci, A., Stornaiuolo, A., D'Apice, M. R., Nigro, V. and Boncinelli, E. (1993). "A vertebrate gene related to *orthodenticle* contains a homeodomain of the bicoid class and demarcates anterior neuroectoderm in the gastrulating mouse embryo." *Embo J* **12**(7): 2735-2747.

Simeone, A., D'Apice, M. R., Nigro, V., Casanova, J., Graziani, F., Acampora, D. and Avantaggiato, V. (1994). "*Orthopedia*, a novel homeobox-containing gene expressed in the developing CNS of both mouse and *Drosophila*." *Neuron* **13**(1): 83-101.

Simeone, A. (2000). "Positioning the isthmic organizer where *Otx2* and *Gbx2* meet." *Trends Genet* **16**(6): 237-240.

Simpson-Brose, M., Treisman, J. and Desplan, C. (1994). "Synergy between the hunchback and bicoid morphogens is required for anterior patterning in *Drosophila*." *Cell* **78**(5): 855-865.

Smith, K. M., Gee, L., Blitz, I. L. and Bode, H. R. (1999). "*CnOtx,* a member of the *Otx* gene family, has a role in cell movement in hydra." *Dev Biol* **212**(2): 392-404.

Stark, D. R., Gates, P. B., Brockes, J. P. and Ferretti, P. (1998). "*Hedgehog* family member is expressed throughout regenerating and developing limbs." *Dev Dyn* **212**(3): 352-363.

Stauber, M., Prell, A. and Schmidt-Ott, U. (2002). "A single *Hox3* gene with composite *bicoid* and *zerknullt* expression characteristics in non-Cyclorrhaphan flies." *Proc Natl Acad Sci U S A* **99**(1): 274-279.

Stornaiuolo, A., Bayascas, J. R., Salo, E. and Boncinelli, E. (1998). "A homeobox gene of the *orthodenticle* family is involved in antero-posterior patterning of regenerating planarians." *Int J Dev Biol* **42**(8): 1153-1158.

Subramanian, V., Meyer, B. and Evans, G. S. (1998). "The murine *Cdx1* gene product localises to the proliferative compartment in the developing and regenerating intestinal epithelium." *Differentiation* **64**(1): 11-18.

Trinh, K. Y., Jin, T. and Drucker, D. J. (1999). "Identification of domains mediating transcriptional activation and cytoplasmic export in the caudal homeobox protein Cdx-3." *J Biol Chem* **274**(9): 6011-6019.

Ueki, T., Kuratani, S., Hirano, S. and Aizawa, S. (1998). "*Otx* cognates in a lamprey, *Lampetra japonica*." *Dev Genes Evol* **208**(4): 223-228.

Umesono, Y., Watanabe, K. and Agata, K. (1999). "Distinct structural domains in the planarian brain defined by the expression of evolutionarily conserved homeobox genes." *Dev Genes Evol* **209**(1): 31-39.

Van den Akker, E., Forlani, S., Chawengsaksophak, K., de Graaff, W., Beck, F., Meyer, B. I. and Deschamps, J. (2002). "Cdx1 and Cdx2 have overlapping functions in anteroposterior patterning and posterior axis elongation." *Development* **129**(9): 2181-2193.

Vandendries, E. R., Johnson, D. and Reinke, R. (1996). "*orthodenticle* is required for photoreceptor cell development in the *Drosophila* eye." *Dev Biol* **173**(1): 243-255.

Wada, S. and Saiga, H. (1999). "Vegetal cell fate specification and anterior neuroectoderm formation by *Hroth*, the ascidian homologue of *orthodenticle/otx*." *Mech Dev* **82**(1-2): 67-77.

Wada, S., Sudou, N. and Saiga, H. (2004). "Roles of *Hroth*, the ascidian *otx* gene, in the differentiation of the brain (sensory vesicle) and anterior trunk epidermis in the larval development of *Halocynthia roretzi*." *Mech Dev* **121**(5): 463-474.

Wang, W. and Lufkin, T. (2000). "The murine *Otp* homeobox gene plays an essential role in the specification of neuronal cell lineages in the developing hypothalamus." *Dev Biol* **227**(2): 432-449.

Wieschaus, E., Perrimon, N. and Finkelstein, R. (1992). "orthodenticle activity is required for the development of medial structures in the larval and adult epidermis of *Drosophila*." *Development* **115**(3): 801-811.

Williams, N. A. and Holland, P. W. (1998). "Gene and domain duplication in the chordate *Otx* gene family: insights from amphioxus *Otx*." *Mol Biol Evol* **15**(5): 600-607.

Yanze, N., Groger, H., Muller, P. and Schmid, V. (1999). "Reversible inactivation of cell-type-specific regulatory and structural genes in migrating isolated striated muscle cells of jellyfish." *Dev Biol* **213**(1): 194-201.

Younossi-Hartenstein, A., Green, P., Liaw, G. J., Rudolph, K., Lengyel, J. and Hartenstein, V. (1997). "Control of early neurogenesis of the *Drosophila* brain by the head gap genes *tll*, *otd*, *ems*, and *btd*." *Dev Biol* **182**(2): 270-283.

CHAPTER III

Involvement of some evolutionarily conserved key players of the "Retinal Determination Genetic Network" in the development, maintenance and regeneration of the eyes of *L.sanguineus*

1. Introduction

1.1 The perception of light

The earth is continuously supplied with electromagnetic radiations from the sun. These radiations represent the primary energy source at the basis of life in most ecosystems, with the exception of few deepwater communities. The light energy is carried in small units, called photons, which have properties of both, waves and particles. Depending on its wavelength, each photon has some energy content: the smaller the wavelength of the photon is, the more energy it carries. Light has probably been one of the most important selective forces during evolution (Fernald, 2000). The capacity to detect light provides valuable information about the illumination or the presence of approaching predator for example. In addition, the information received by an organism about the day/ night cycle allows the creation of a circadian rhythm. Light has influences on organism's photosynthesis, movement and behavior. In comparison to the broad spectrum of energy produced by sunlight, the photoreceptors have evolved a quite narrow range of wavelength detection. This actually reflects the fact that life has originated in an aquatic environment, which strongly filters the light (Fernald, 1988). As the light is strongly filtered in an aquatic environment, the essential selective pressure on early organisms was for light perception within a limited range of wavelengths. During evolution, many animals have moved onto the land and, consequently, have been exposed to a broader spectrum of radiations from the sun. Some organisms, such as insects, some fish and birds have evolved additional receptors for UV light detection (Kevan *et al.*, 2001). However, most of the terrestrial animal eyes remain limited to the detection of the same narrow range of wavelengths than the one detected by early aquatic organisms. The visible light is characterized by medium energy photons, which are absorbed by carotenoid pigments in plants and animals. The visual information can be directly transferred to an effector organ or processed by a brain. For example, in the unicellular algae *Euglena*, the visual information acquired by a photosensitive area, called the stigma, is directly transferred to the locomotory flagella. *Euglena* swims in a helix path so that the stigma-photoreceptor sweeps through a circle. If more light comes from one region than from the others, *Euglena* will turn to it by changing the direction of its flagellar beat (Lebert and Hader, 1997).

The perception of light is a widespread sensory ability: even simple organisms, such as bacteria possess light-sensitive receptor. Their receptor, a G-protein coupled receptor, converts the energy from the light into an electrochemical gradient. Some single-celled dinoflagellates, like *Erythropsis* and *Warnovia*, present already highly elaborated structures for the perception of light (Greuet, 1965). Interestingly, dinoflagellates are common symbionts in cnidaria. Hence, the genes required for photoreception could have been transferred from dinoflagellates to cnidaria. This has led to the tempting hypothesis that the metazoan eye could have originated from symbionts, like chloroplasts, as proposed by the "Russian doll model" of Gehring (Gehring, 2005). However, simple photoreceptors have evolved to the great complexity of image-forming organs, such as the vertebrate lens eye or the arthropod compound eye, only in higher animals. A tremendous variety of eye-types has been generated during the course of evolution: from simple eyespots to more elaborated eyecups, mirror eyes, pin-hole eyes, compound eyes and the different camera-type eyes of cephalopods and vertebrates. All these eyes are morphologically and physiologically very different (Fig. 3.1).

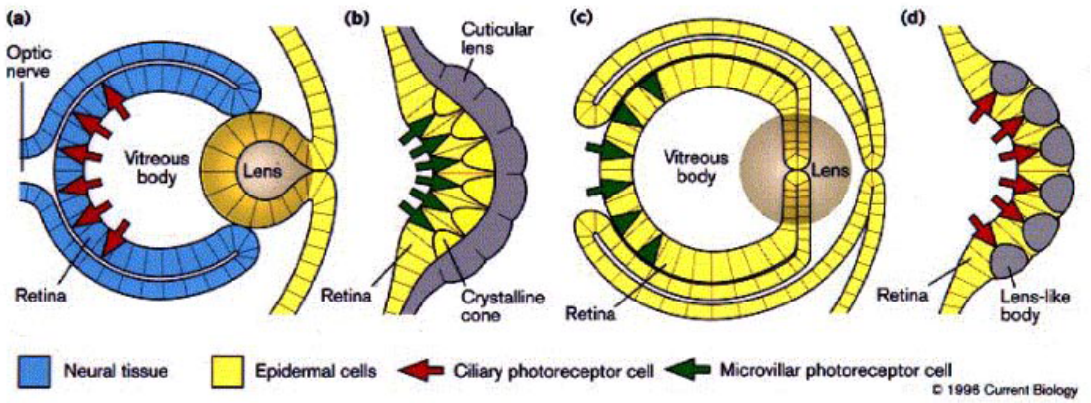


Fig. 3.1 Simplified illustrations of the building-plans of four types of eye (Nilsson, 1996)
a: vertebrate eye; b: arthropod compound eye; c: cephalopod lens-eye; d: compound eye in polychaete tube-worms and arcoid clams.

1.2 Origin of the eyes

Based on anatomical and embryological studies, it has been proposed that the eyes of the various animal phyla have evolved independently, at least 40 to 60 times (Salvini-Plawen and Mayr, 1961). Since then, the polyphyletic origin of the eyes has been considered as a dogma, found in all biological textbooks. However, one should notice that all the three major eye-types, the camera eye, the compound eye and the mirror eye, are found within a single phylogenetic class, the Bivalvia. Indeed, the camera-type eye, with a single lens, is present in the heart shell *Cardium*. The mirror-type eye, with a lens and a reflecting mirror, is found in the scallop *Pecten*. And finally, the compound-type eye, consisting of 10 to 80 ommatidia each, is observed in *Arca noae*. Hence, the idea of an independent origin of these three eye-types within the same class of mollusk is not really likely.

Furthermore, the theory of a polyphyletic origin of the eyes has been greatly challenged by the discovery that homologous genes, the *Pax-6* genes, initiate eye formation in almost all bilaterian animals. Indeed, *Pax-6* is essential for eye differentiation in both, *Drosophila* and vertebrates (Hill *et al.*, 1991; Quiring *et al.*, 1994). Moreover, it has been shown to be sufficient for the induction of ectopic eyes in certain body regions of *Drosophila* and vertebrates (Halder *et al.*, 1995; Chow *et al.*, 1999; Onuma *et al.*, 2002). This highlights the conserved capacity of *Pax-6* to act as a "master control gene" of eye development (Gehring and Ikeo, 1999). Subsequently, a wealth of additional similarities has been found between vertebrate and invertebrate eyes, at the molecular level (Wawersik and Maas, 2000; Arendt and Wittbrodt, 2001; Donner and Maas, 2004). Consequently, it seems that evolutionarily conserved genes are deployed, as "shared building blocks", to generate very divergent morphological eyes. This apparent paradox is actually, not unique to the eyes, and seems rather to be an emerging generality of the evolution of development (Arthur, 2002). Based on these molecular similarities, Gehring and Ikeo have proposed a possible monophyletic origin of all different eye-types (Gehring and Ikeo, 1999): this evolutionary model suggests that a common metazoan ancestor had a prototypic eye, composed of a single pigment cell and a single photoreceptor cell and lacking any lens or other refractive body, already under the control of the conserved genetic "building blocks". In accordance with Darwin's theory, such primitive eyes indeed exist in some adult Planaria and also in *Platynereis* larvae (Arendt *et al.*, 2002;

Gehring, 2004). The prototypic eye could have given rise to the wide spectrum of eyes existing today through the evolution of developmental pathways by an "intercalary evolution" mechanism (Gehring and Ikeo, 1999). The developmental pathway of the prototypic eye could have been composed of *Pax-6* at the top and of structural genes, such as the *opsin* genes, at the bottom. During evolution, new genes would have been intercalated between the top and the bottom of the cascade. Hence, the "intercalary evolution" model suggests that the addition of some new developmental genetic processes to the conserved ones could have resulted in morphological changes (Gehring and Ikeo, 1999). The mechanism of "replication and divergence", as a central concept in molecular evolution (Raff, 1996), is consistent with the monophyletic origin hypothesis and the "intercalary evolution" model (Oakley, 2003). It proposes that replication of eye structures, at new sites on a organism, under new selective constraints, would have allowed the addition of new genes or network in its developmental program. Subsequently, divergence, loss or maintenance of eye structures would have given rise to the wide spectrum of morphologically different eyes, which could nevertheless be traced back to a single prototypic eye (Oakley, 2003). By contrast, some authors have argued for an independent origin of the eyes and for a subsequent recruitment or co-option of common genes for their development (Fernald, 2000; Nilsson, 2004).

A novel evo-devo approach of cell type comparison, the comparative molecular cell biology, might help to gain new insights into the eye evolution (Arendt, 2003). Based on the expression of developmental genes and also, on the effector genes that are responsible for the cell's function, this approach allows the identification of homologous cell-types. These cells have evolved from a same precursor cell present in the last common ancestor of the compared groups. Using this comparative molecular cell biology approach, it has been proposed that today's photoreceptor cells have evolved from a "photoreceptor cell precursor". This photoreceptor cell precursor was already using arrestin and rhodopsin kinase for quenching the light signal (see Fig. 3.6). In addition, it has been proposed that at least, Pax, Otx, and Six transcription factors are representing the ancestral combinatorial code for photoreceptor cell fate specification and differentiation (Arendt, 2003).

1.3 Similarities and differences in various eye-types

1.3.1 Structure and developmental mode of the *Drosophila* compound eye *vs* the vertebrate eye.

Compound eyes are mostly, but not exclusively, found in arthropods. The *Drosophila* compound eye forms a regular hexagonal array, consisting of approximately 800 single units, the ommatidia. Each ommatidium contains eight photoreceptor cells, seven pigment cells, a mechanosensory bristle and four cone cells, which are secreting a lens. The photoreceptors project into the lamina and medulla, two optic ganglia of the brain.

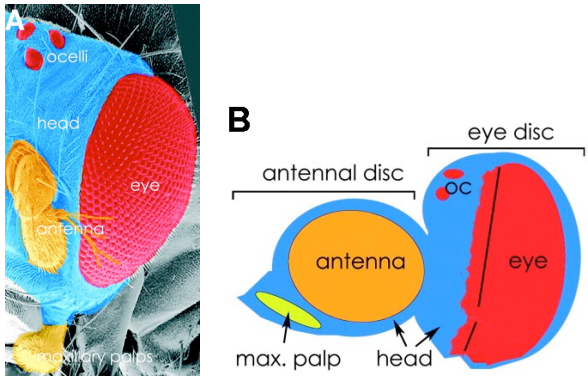


Fig. 3.2 The adult head of *Drosophila* develops from a pair of composite discs called the eye-antennal discs (Dominguez and Casares, 2005)
A: a scanning electron photomicrograph of a wild-type female adult head. The discrete organs of the adult head are marked in the same color code as the corresponding organ-forming primordia in the eye-antennal imaginal disc in B.

During metamorphosis, different *Drosophila* adult epidermal structures and genitalia develop out of sac-like clusters of primordial cells, the imaginal discs. Most of the head capsule and the major cephalic sensory organs, like the eyes, the antennae, the maxillary palps, and the ocelli, derive from the eye-antennal discs (Fig. 3.2). This monolayer epithelium is formed by an involution from the head ectoderm during the embryonic stages. The differentiation of the eye disc is progressive, moving as a wave from posterior to anterior across the eye imaginal disc, in the third larval stage. The wave is visible in the imaginal disc as an indentation, also called the morphogenetic furrow. Anterior to the furrow, where proliferative cell division takes place, cells are un-patterned, while cells behind the furrow are organized into clusters, progressively differentiating into functional ommatidia.

The vertebrate eye is composed of an inverted neural retina and a pigmented retina, transparent and refractive cornea and lens, and a contractile iris, which controls the amount of light admitted through the lens. All together, the cornea, lens and iris form a focusing light system that project the light onto the neuroretina.

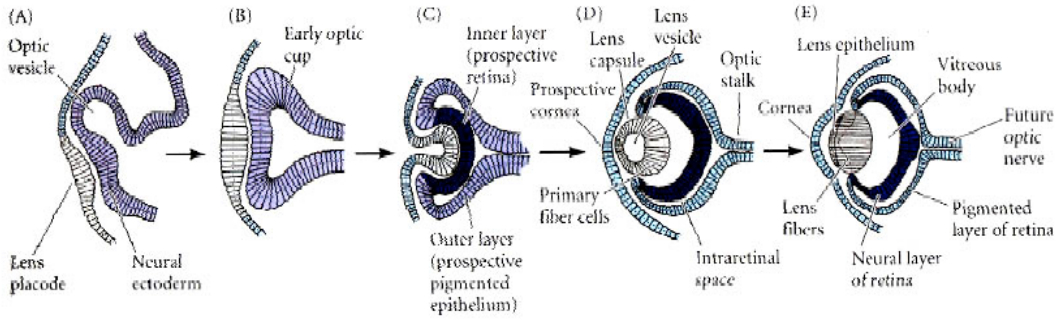


Fig. 3.3 Development of the vertebrate eye (Cvekl and Piatigorsky, 1996)
A: the optic vesicle, a protuberance of the forebrain, extends towards the surface ectoderm; the lens appears as a local thickening of the surface ectoderm. B: the optic vesicle becomes an optic cup. C: the lens placode invaginates and forms the lens vesicle. The two layers of the optic cup, the neuroretina and the retinal pigmented epithelium, are distinguishable. D: the lens vesicle induces the development of the cornea. D: Cross-section of a developed vertebrate eye.

The development of the vertebrate eye starts at gastrulation, where the involuting endoderm and mesoderm interact with the adjacent ectoderm (Fig. 3.3). Later, after neurulation, the neural tube evaginates laterally to form the two optic vesicles, which are connected to the diencephalon (forebrain) by the optic stalks. The optic vesicles contact the surface ectoderm, which, in turn, thickens into the lens placode. Through mutual interaction of the optic vesicle with the lens placode of the overlaying head ectoderm, the optic vesicle invaginates and forms a double-layered optic cup. The outer layer will give rise to the retinal pigmented epithelium, while the inner layer differentiates into the neural retina. The lens vesicle induces the development of the transparent cornea. The soluble crystallin proteins accumulate at a high concentration within corneal cells. This leads to the acquisition of the cornea specific refractive properties (Cvekl and Piatigorsky, 1996).

Drosophila and vertebrate eyes are clearly very different at the structural level. Strikingly, even eyes of the same type can exhibit fundamental differences: the vertebrate retina is inversed, with photoreceptors facing the back of the eye, whereas the cephalopod retina is eversed, with photoreceptors facing the front of the eye. In addition, the various eye-types display very different developmental modes (Nilsson, 1996). However the developmental program of both *Drosophila* and vertebrate eyes might be more related than expected at a first sight. Indeed, both eye types are patterned by a conserved mechanism, a morphogenetic wave (Fig. 3.4): in *Drosophila*, the Hedgehog signaling molecule drives the wave of neurogenesis, represented by the morphogenetic furrow (Heberlein *et al.*, 1995). Similarly, there is evidence that the vertebrate *hedgehog* homologue, *sonic hedgeghog*, patterns the zebrafish retina (Neumann and Nuesslein-Volhard, 2000).

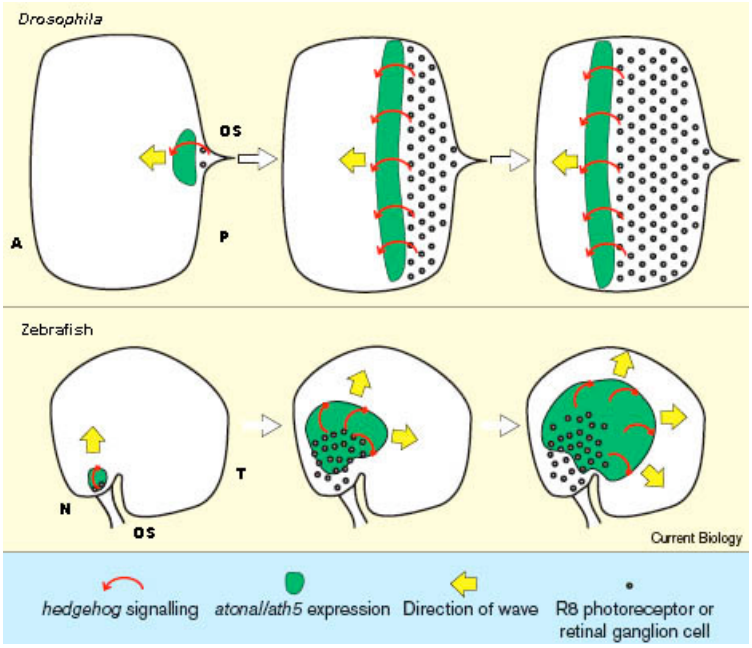


Fig. 3.4 The Mexican wave in eye development (Jarman, 2000)
Schematic drawing of three different time points in the progression of the initial neurogenesis in the *Drosophila* eye imaginal disc and in the zebrafish inner optic cup, below. In both cases, atonal and its vertebrate homologue, *ath5*, are involved in determining the first neurons. The short-range hedgehog signaling molecule drives the wave of neurogenesis in both cases. In *Drosophila*, it is via the activation of *atonal*, while it seems likely to be via *ath5* activation in zebrafish. A: anterior; P: posterior; N: nasal; T: temporal.

1.3.2 The ultrastructure of photoreceptor cells

Photoreceptors are neurons specialized for photoreception. The ones, involved in vision, are characterized by an enlargement of their membrane surface. This feature allows the storage of a large amount of rhodopsin, which is the photoreception molecule used in all animal visual photoreceptors (Salvini-Plawen and Mayr, 1977). According to this ultrastructural feature, the photoreceptors are divided into two classes: the rhabdomeric photoreceptors and the ciliary photoreceptors, in which the apical cell membrane or the ciliary membrane is enlarged, respectively (Fig. 3.5).

Most of invertebrate eyes possess rhabdomeric photoreceptors, in which the apical cell surface folds into microvilli. In contrast, the rods and cones of the vertebrate retina and the photoreceptors of the vertebrate pineal eye, a light-sensitive structure located in the diencephalon, are ciliary photoreceptors. It has been initially proposed that rhabdomeric photoreceptors are specific for invertebrate eyes, while ciliary photoreceptors were characteristic of vertebrate eyes. However, it turns out that both types of photoreceptors co-exits in Lophotrochozoa, Ecdysozoa and Deuterostomia (Arendt and Wittbrodt, 2001). Furthermore, many organisms possess both types, sometimes even within a single eye, as in the scallop *Pecten*. Therefore, it has been proposed that both type of photoreceptors were already present in the common ancestor of bilaterian animals (Arendt and Wittbrodt, 2001).

1.3.3 Photopigments

Light-sensitive pigments are found in almost all living organisms, from bacteria to humans. The photopigments of unicellular algae are flavoproteins, including cryptochromes, phototropins and photo-activated adenyl cyclases (Iseki *et al.*, 2002). Cryptochromes are also found in insects and vertebrates, where they control the biological clock and consequently maintain proper circadian rhythms (Van der Horst *et al.*, 1999; Krishnan *et al.*, 2001). In contrast to the flavoproteins, other photopigments, known as the retinal-binding opsins, are linked to a transduction cascade, which generates a receptor potential. The opsins are G-protein coupled receptors (GPCRs). Hence, they possess a seven-transmembrane structure but are distinct from other GPCRs by having a lysine residue, which is a retinal binding-site, in the seventh helix (Terakita, 2005). Even if animal photopigments involved in

vision are exclusively of the opsin family, some types of ospin are also involved in extra-ocular photoreception (Menaker, 2003). Sequence comparisons have indicated that ciliary ospins (c-opsins) and rhabdomeric opsins (r-opsins) are highly divergent molecules. The c-opsins are closer to the retinochromes than to the r-opsins, while the r-opsins are more related to vertebrate melanopsins than to the c-opsins (Arendt and Wittbrodt, 2001; Terakita, 2005).

1.3.4 The extremely sensitive rhodopsin molecule

A photosensitive rhodopsin molecule is capable of sensing a single photon and of turning the physical energy contained in this single light quantum into an electrochemical signal. A rhodopsin molecule is composed of a protein moiety, an opsin, and a non-protein moiety, the chromophore retinal, which is a derivative of vitamin A and is covalently bound to the opsin (Terakita, 2005). The absorption of a single photon triggers the isomerization of the *11-cis* retinal into an *all-trans* retinal. Hence, the rhodospin molecule is photoactivated, also defined as rhodopsin in its meta-state. The photoactivated rhodospin activates a heterotrimeric GTP-binding protein (G-protein), which, in turn, activates intracellular messengers.

1.3.4.1 The phototransduction cascade

The phototransduction machinery differs between the ciliary and the rhabdomeric photoreceptors (Fig. 3.5) (Arendt and Wittbrodt, 2001). In fact, each type of photoreceptor employs its own subgroup of non-orthologous G-proteins. In addition, the second intracellular messengers also differ in the phototransduction cascades occurring in rhabdomeric and in ciliary photoreceptors. In rhabomeric photoreceptors, the G-protein activates a phospholipase enzyme (PLC), which converts phosphatidyl inositol diphosphate (PIP_2) into inositol triphosphate (IP_3). In ciliary photoreceptors, the G-protein activates phosphodiesterase (PDE), which converts cyclic guanosine monophosphate ($_C$GMP) into guanosine monophosphate (GMP).

The quenching of the phototransduction cascade requires rhodopsin kinases and arrestins. Theses molecules are also non-orthologous in ciliary versus rhabdomeric photoreceptors. Finally, the transduction cascade eventually leads to a depolarization or a hyperpolarization of the membrane potential in rhabdomeric and in ciliary photoreceptors, respectively (Arendt and Wittbrodt, 2001).

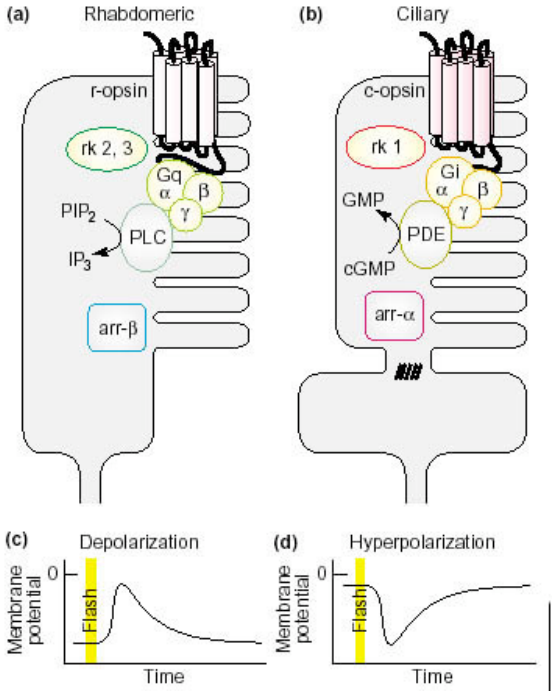


Fig. 3.5 The two types of photoreceptors: the rhadomeric type and the ciliary type (Nilsson, 2004)
The photopigments (r-opsin and c-opsin) are activated by the absorption of light. Then, the opsins activate a G-protein, which is composed of three subunits (α, β, γ). In turn, the G-protein activates an enzyme, a PLC or a PDE in rhabdomeric and ciliary photoreceptors, respectively. This finally leads to a depolarization of the rhabdomeric photoreceptor or a hyperpolarization of the ciliary photoreceptor, respectively. rk: rhodopsin kinase, arr: arrestin.

In summary, the light detection and phototransduction systems of the rhabomeric photoreceptors use a r-opsin, a Gq-protein, a second messenger system based on PLC, RK-2/3 rhodospin kinases related and β-arrestin, and eventually lead to membrane potential depolarization. In contrast, the light detection and phototransduction systems of the ciliary photoreceptors use a c-opsin, a Gi or Go-protein, a second messenger system based on PDE, RK-1/4/5/6-related rhodospin kinases and α-arrestin, and eventually lead to a membrane potential hyperpolarization.

1.3.4.2 Rhodopsin regeneration after photobleaching

Once a rhodopsin molecule has been activated by one photon, it is rapidly deactivated and regenerated. Differences between the visual pigment recovery path of rhabdomeric and ciliary photoreceptors exist (Fig. 3.6). Every bleached rhodospin molecule is rapidly inactivated through phosphorylation by a rhodospin kinase and also, through binding to arrestin.

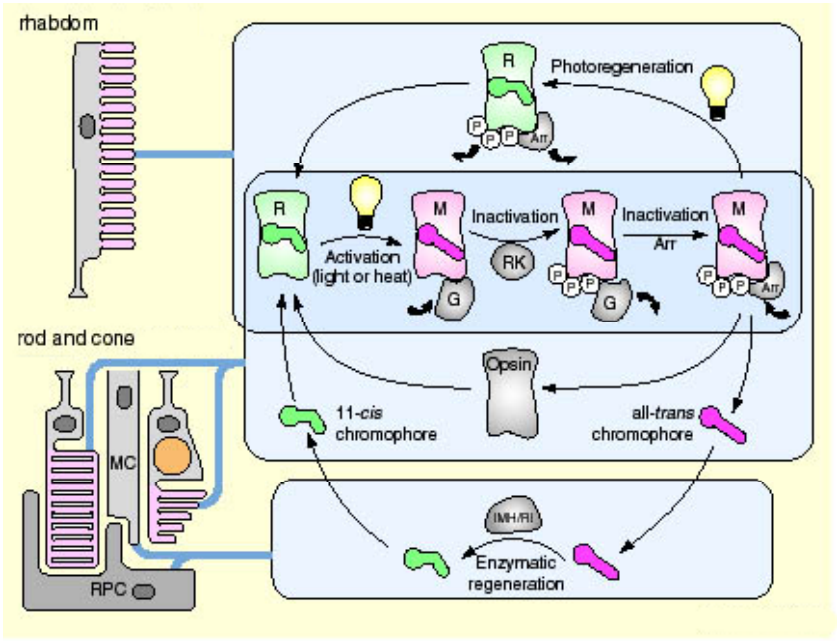


Fig. 3.6 The different recovery paths of the rhabdomeric and ciliary visual pigments (Nilsson, 2004)
Due to the absorption of a photon, the retinal isomerizes from an 11-cis to an all-trans configuration. This turns the rhodopsin (R) molecule into its meta-state (M), leading to catalytic activation of a G-protein (G) and the phototransduction cascade. The activated rhodopsin is rapidly inactivated through phosphorylation (P) by rhodopsin kinase and through binding to arrestin (Arr). In ciliary photoreceptors, the retinal is transferred to other cells and is enzymatically reconverted to an 11-cis configuration. In vertebrate rods retinal pigment cells (RPCs) reconvert the retinal, whereas Müller cells (MCs) serve the same function for vertebrate cones. The enzymatic reaction is done by an isomerohydrolase (IMH) and a retinol isomerase (RI) in RPCs and in MCs, respectively. The 11-cis retinal is then transported back to rods and cones, where it is incorporated into an empty opsin molecule. For the recovery of their visual pigment, the rhabdomeric photoreceptors use a light-driven regeneration mechanism, in which the retinal never dissociates from the opsin.

Differences exist in how the visual pigment is regenerated: in ciliary photoreceptors, the *all-trans* retinal dissociates from the opsin and is transferred to other cells, where it is converted into an *11-cis* configuration by an enzymatic reaction. Then, the *11-cis* retinal is transported back to the ciliary photoreceptors and

associates again with an opsin molecule. Like this, the visual pigment of ciliary photoreceptors is regenerated. In contrast, in rhabdomeric photoreceptors, the retinal never dissociates from the opsin. The *all-trans* retinal is reisomerized into an *11-cis* configuration by a light-driven regeneration process, occurring directly in the rhabdomeric photoreceptor.

1.3.5 Single origin of photoreceptor cells

Traditionally, the evolutionary origins of the two types of photoreceptors have been considered to be different. Rhabdomeric photoreceptors were thought to be specific for invertebrate eyes. Reciprocally, ciliary photoreceptors were thought to be specific for vertebrate eyes. This dichotomy was considered as a real dogma. However, based on anatomical studies only, ciliary-type photoreceptors have been described in some invertebrate brains. It is also known that both, rhabdomeric photoreceptors and ciliary photoreceptors are present in the eyes of the scallop *Pecten*. In addition to anatomical data, recent molecular data have greatly challenged the theory of a diphyletic origin of both photoreceptor types.

Indeed, ciliary photoreceptors, using a c-opsin involved in a ciliary-type phototransducion cascade, have been found in the brain of *Platynereis*, a lophotrochozoan member (Arendt *et al.*, 2004). The development of these brain photoreceptors is controlled by the gene *rx*, which is homologous to the gene that regulates the development of vertebrate ciliary photoreceptors. Hence, based one their developmental control, their structure, their constituents and their signaling machinery, these photoreceptors can truly be identified as “ciliary photoreceptors”. Likewise, findings in neurophysiology and data from other studies argue for the presence, within the vertebrate lineage, of true rhabdomeric photoreceptors, in the form of retinal ganglion cells involved in circadian and pupillary adjustments (Dacey *et al.*, 2005; Panda *et al.*, 2005). These retinal ganglion cells use melanopsin, a r-opsin type, and have a rhabdomeric phototransduction cascade type (Isoldi *et al.*, 2005). On top, homologues of the *atonal* gene are required for the development of the rhabdomeric photoreceptors and retinal ganglion cells in flies and in vertebrates, respectively.

Thus, true homologous rhabdomeric photoreceptor cells, usually considered as invertebrate characteristics, are found also in vertebrates. Reciprocally, true homologous ciliary photoreceptor cells, usually considered as vertebrate characteristics, are present in invertebrates, as well. Combined all together, these data strongly support an ancient and singular origin of both photoreceptor cell types (Plachetzki *et al.*, 2005).

1.4 The retinal determination gene network: the RDGN

The specification of the eye field of various organisms requires the expression of homologous members of the retinal determination gene network (the RDGN) (Fig. 3.8). The RDGN plays a pivotal role for eye development by integration multiple signaling pathways. It affects and, reciprocally, is affected by several signaling pathways, in a context-specific manner (Fig. 3.7).

For example, Dpp, Hh and Wg signaling are known to regulate the expression of the RDGN members (Hazelett *et al.*, 1998; Baonza *et al.*, 2002). In addition, a complex interface between EGFR signaling and the RDGN is suggested by multiple genetic interactions (Silver and Rebay, 2005). Members of the Pax-6, Eyes absent (Eya), Six (Six) and Dachshund (Dach) families are key members of the RDGN. They do not function in a linear pathway, but rather in a complicated network involving mutual interactions and regulatory feedback loops. In *Drosophila*, seven "eye-specification genes" are part of the network that initiates eye development (Fig. 3.7): the *Pax-6* homologues: eyeless (*ey*) and twin of eyeless (*toy), sine oculis* (*so*), *optix, eyes absent* (*eya*), *dachshund* (*dac*) and *eye gone* (*eyg*) (Halder *et al.*, 1998; Niimi *et al.*, 1999; Bui *et al.*, 2000; Seimiya and Gehring, 2000). The accepted model for *Drosophila* eye induction is that *ey* induces the initial expression of *so* and *eya* that, in turn, regulate the activity of all four core genes of the RDGN: *ey*, *so* and *eya,* and *dac*. Removal of any of these genes in the eye primordium results in a severe reduction or complete loss of the adult compound eye, while ectopic expression of these genes (with the notable exception of *so*) leads to the induction of ectopic eye tissue (Halder *et al.*, 1995; Bonini *et al.*, 1997; Shen and Mardon, 1997; Czerny *et al.*, 1999; Seimiya and Gehring, 2000). Most combinations of ectopic expressions of these RDGN members result in a synergistic induction of ectopic eye tissue formation.

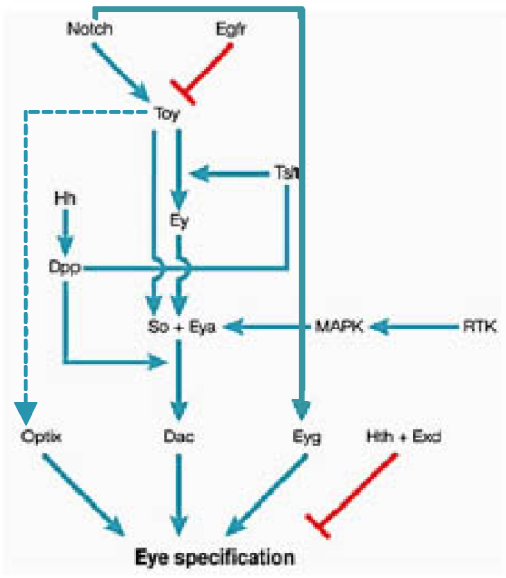


Fig. 3.7 Eye specification genes in *Drosophila* (modified after Kumar, 2001)
Several nuclear factors, patterning pathways and signaling cascades orchestrate, in a complicated network, the specification and development of the *Drosophila* eye. The arrows show the direction of the relationship: blue ones indicate activation, whereas red ones indicate inhibition. Dac, Dachshund; Dpp, Decapentaplegic; Egfr, Epidermal growth factor receptor; Exd, Extradenticle; Ey, Eyeless; Eya, Eyes absent; Eyg, Eye gone; Hh, Hedgehog; Hth, Homothorax; MAPK, Mitogenactivated protein kinase; RTK, receptor tyrosine kinase; So, Sine oculis; Toy, Twin of eyeless; Tsh, Teashirt.

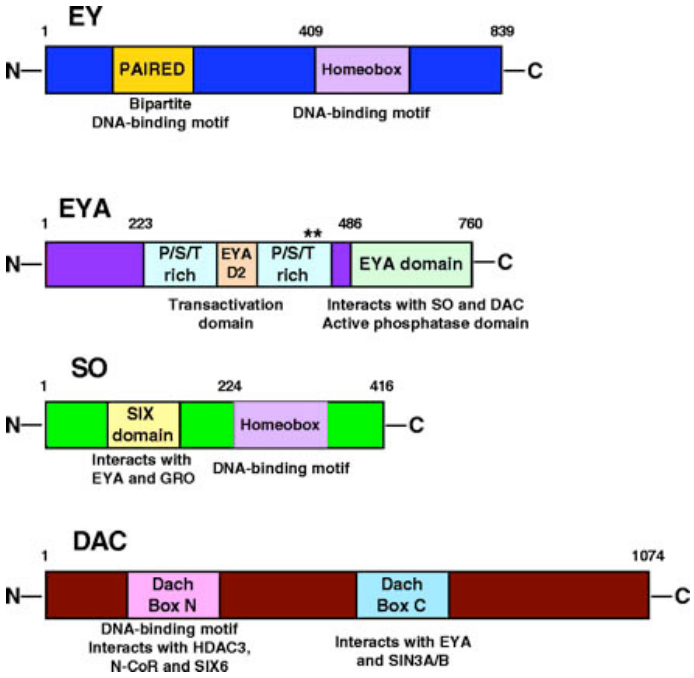


Fig. 3.8 Domain structures of the *Drosophila* RDGN key members (Silver and Rebay, 2005)
Numbers represent amino acid number. EY: eyeless; EYA: eyes absent; P/ S/ T rich: proline, serine and threonine rich region, EYA D2: EYA domain 2; **: conserved MAPK phosphorylation sites in EYA; SO: sine oculis; GRO: groucho; DAC: dachshund; HDAC3: histone deacetylase 3; N-CoR: nuclear co-repressor.

1.4.1 The Pax-6 transcription factor

Pax-6 belongs to the superfamily of *Pax* genes, which encode evolutionarily highly conserved transcription factors. These transcription factors are characterized by the presence of a 128 aa DNA-binding domain: the paired domain (PD), which is a bipartite DNA-binding motif (Treisman *et al.*, 1991). It binds DNA alone or cooperatively with other DNA-binding domains, such as the HD. The PD has two independent subdomains: a N-term domain: PAI and a C-term domain: RED, separated by a linker region. Both subdomains consist of three alpha helices, arranged in a helix-turn-helix configuration (Xu *et al.*, 1995) and both can recognize independent binding sites (Czerny *et al.*, 1993). However, when both PD subdomains are present, the PAI binding site is preferred for binding to the DNA. Originally found in the *paired* gene of *Drosophila* (Bopp *et al.*, 1986), the paired box has now been identified in all metazoan phyla, from placozoans (Hadrys *et al.*, 2005) and cnidarians (Sun *et al.*, 1997) to humans (Burri *et al.*, 1989). In addition to the paired domain, other motifs can be found in Pax proteins: an octapeptide, which has been shown to possess transcriptional inhibitory activity in the *Pax 2/5/8* family (Eberhard *et al.*, 2000), and a complete or partial HD. The paired-type HD is characterized by the presence of a serine at the crucial position 50. The various members of the Pax superfamily appear to result from a combinatorial rearrangement of the PD, the HD (complete or partial) and the octapeptide (Fig. 3.9). This is actually a very elegant example of what François Jacob called "evolutionary tinkering" (Jacob, 1977). In mammals, nine *Pax* genes have been identified, denoted *Pax-1* to *Pax-9*. Based on their genomic structure, sequence similarity and conserved function, *Pax* genes have been grouped into four subfamilies: (1) *Pax-1* and *Pax-9*; (2) *Pax-3* and *Pax-7*; (3) *Pax-4* and *Pax-6*; (4) *Pax-2, Pax-5* and *Pax-8* (Fig. 3.9). *Pax* genes are key regulators of numerous developmental processes, in particular neurogenesis and myogenesis. Some of them have also been implicated in regenerative processes, as well as oncogenesis. Their general role might be the interpretation of "positional information", via signal transduction and cell proliferation (Chi and Epstein, 2002; Pichaud and Desplan, 2002).

	Structural domains			
	PD	OP	HD1	HD2/3
PAX1	X	X		
PAX9	X	X		
PAX2	X	X	X	
PAX5	X	X	X	
PAX8	X	X	X	
PAX4	X		X	X
PAX6	X		X	X
PAX3	X	X	X	X
PAX7	X	X	X	X

Fig. 3.9 Structural domains present in the different Pax proteins
PD: paired domain; OP: octapeptide; HD1: the first helix of the homeodomain and HD2/3: the helix-turn-helix motif of the homeodomain. X denotes the presence of the domain.

Pax-6 genes were first identified in mammals (Ton *et al.*, 1991; Walther and Gruss, 1991). In mice and humans, eye defects are observed in *Pax-6* heterozygote mutants, in the *Small eye* mutants and *Aniridia* patients, respectively. Homozygous *Pax-6* mutation is lethal to mouse embryos as they, not only, lack the eye and the nose but also, exhibit severe brain damage. In *Drosophila*, two *Pax-6* genes, *eyeless* (*ey*) (Quiring *et al.*, 1994) and *twin of eyeless* (*toy*) (Czerny *et al.*, 1999), have been identified. They probably arose by gene duplication in the arthropod lineage. Based on sequence comparisons and DNA-binding specificities, *toy* seems to be closer to other *Pax-6* homologues than *ey*. Hypomorphic *eyeless* mutants result in partial to complete loss of the compound eyes. Both, *ey* and *toy*, are required for *Drosophila* eye formation. *toy* acts upstream of *ey* and induces its expression in the eye anlage. The PD of Ey is sufficient for its function in eye development. Indeed, it has been shown that the HD of Ey is dispensable for *Drosophila* eye development (Punzo *et al.*, 2001).

Pax-6 is essential for eye development and is required for the expression of the downstream RDGN members. Furthermore, *Pax-6* is sufficient to induce ectopic eyes, upon misexpression, in both, flies and vertebrates, by initiating a cascade of at least 2000 genes required for eye morphogenesis (Halder *et al.*, 1995; Chow *et al.*, 1999; Onuma *et al.*, 2002). Hence, *Pax-6* has been proposed to be at the top of the transcriptional cascade of the RDGN.

Based on expression patterns and mutant phenotypes, *Pax-6* is a crucial regulator of eye development, in the vast majority of animals. In addition, it also plays important roles in the development of the CNS and in the development of some endocrine glands, such as the pineal gland and the pituitary gland, in both, invertebrates and vertebrates. The evolutionary conserved requirement of *Pax-6* for CNS morphogenesis explains why, in contrast to the *opsin* (Iwabe *et al.*, 1996; *C.elegans* Sequencing consortium, 1998), this gene has been conserved intact in organisms, such as the nematode *C.elegans* (Chisholm and Horvitz, 1995), which have secondarily lost their eyes.

1.4.2 The Eya transcription factors

Eya family proteins are characterized by a large highly conserved C-term domain of 275 aa, the Eya domain (ED) (Fig. 3.8). *Drosophila* has a single *eya* gene (Bonini *et al.*, 1993), while the vertebrates possess four homologues, *Eya-1* to *Eya-4* (Xu *et al.*, 1997). Studies from the fly *eya* and the vertebrate homologues indicate that they play important roles in cell survival and differentiation, especially during tissue specification.

The function of the Eya homologues is greatly conserved. Indeed, vertebrate *Eya* genes can rescue the “eyeless” phenotype of *Drosophila eya* mutants (Bonini *et al.*, 1997; Bui *et al.*, 2000). The ED mediates protein-protein interactions with other RDGN members in flies and vertebrates: notably with the Six-domain of the Six proteins (Pignoni *et al.*, 1997; Ohto *et al.*, 1999) an d with the DachBox-C domain of Dach proteins (Nilsson, 1996; Chen *et al.*, 1997; Heanue *et al.*, 1999). Eya acts as transcriptional co-activator, which is recruited to the DNA of target genes via its interaction with Six proteins (Silver *et al.*, 2003). This transcriptional co-activator

function is dependent on N-terminal domains of Eya (Silver *et al.*, 2003). These N-terminal domains are composed of a tyrosine rich domain, called the Eya domain 2, which is embedded within a proline/ serine/ threonine-rich region. In addition to its transcriptional co-activator function, Eya exhibits protein phosphatase activity mediated by its ED, which possesses a catalytic motif from the haloacid dehalogenase enzyme family (Rayapureddi *et al.*, 2003; Tootle *et al.*, 2003). This is actually the first description of a nuclear transcriptional co-activator with intrinsic phosphatase activity. The process of phosphorylation/ dephosphorylation is known to be extremely important for the modulation of transcription factor activity. Therefore, the discovery of a transcription factor with an intrinsic phosphatase activity argues for a novel strategy for fine-tuning regulation of transcription (Rebay *et al.*, 2005). Eya can dephosphorylate the RNA polymerase II and itself, in vitro. In addition, mutations disrupting the active site, responsible for the Eya phosphatase activity, within the ED, compromise the ability of Eya to promote *Drosophila* eye specification and development (Tootle *et al.*, 2003).

1.4.3 The Six family members

Six genes have been identified throughout bilaterian animals and also in several representatives of basal metazoans, such as sponges, cnidarians and in a ctenophore (Bebenek *et al.*, 2004; Stierwald *et al.*, 2004). These genes encode transcription factors that are involved in numerous developmental processes, such as eye formation, forebrain development and myogenesis, and play important roles in the regulation of cell proliferation (Carl *et al.*, 2002; Li *et al.*, 2002). The Six family members are characterized by the presence of two highly conserved domains: a Six-domain (SD) of 110 to 115 aa and a Six-type HD (Fig. 3.8). The Six-type HD lacks the usual Arg5 and the Gln12 in helix 1 and therefore does not recognize the conserved homeobox binding core sequence TAAT. Both domains, SD and HD, are involved in DNA binding. In addition, the SD also mediates protein-protein interactions and is responsible for nuclear translocation of the EYA members (Kawakami *et al.*, 2000). Based on sequence conservation within HD and SD, the *Six* family is subdivided into three subclasses, each containing one *Drosophila* member and two vertebrate homologues: *so*/ vertebrate *Six1* and *Six2; D-Six4*/ vertebrate *Six4* and *Six5*; *optix*/ vertebrate *Six3* and *Six6* (Seo *et al.*, 1999; Kawakami *et al.*, 2000).

The So/ Six1/ Six2 and DmSix4/ Six4/ Six5 protein subfamilies interact with the ED of Eya proteins (Ohto *et al.*, 1999). Interestingly, the Optix/ Six3/ Six6 protein subfamily does not interact with the ED of Eya proteins and recognizes different target sequences (Ohto *et al.*, 1999). Studies from vertebrates suggest that Six3/ Six6 act as transcriptional repressors, through their interactions with the Groucho (Gro) family of co-repressors in order to achieve proper eye and brain formation (Lopez-Rios *et al.*, 2003). The Six3/ Six6 proteins interact with GRO co-repressors via an Engrailed homology 1 (eh1) motif (Kobayashi *et al.*, 2001). As this eh1 motif is present in the SD of all Six proteins, other Six members are likely to exhibit as well a transcriptional repression function. Indeed, So is able to interact with *Drosophila* Gro (Silver *et al.*, 2003) and a transcriptional repression function has also been revealed for the mouse Six1 (Li *et al.*, 2003). Thus, depending of the context and of specific co-factors, a Six protein might be able to act either as an activator or as a repressor. However, the Six3/ Six6 have so far only been reported as transcriptional repressors.

In *Drosophila*, So is essential for the development of the entire visual system: it has been shown to be required for the development of both, the compound eyes and the ocelli (Cheyette *et al.*, 1994). It is a direct target of both Toy and Ey, and interacts with Eya (Punzo *et al.*, 2002). In mouse, Six1 displays important roles in the formation of numerous organs, such as the nose, the inner ear, the kidney and the pituitary gland (Laclef *et al.*, 2003a; Ozaki *et al.*, 2004). It is known to interact with members of the RDGN, namely Dach2, Pax3 and Eya2, for the proper formation of the somites and its skeletal muscles derivatives (Heanue *et al.*, 1999). This illustrates that members of the RDGN can be viewed as “building blocks”, which have been reused during evolution to trigger development of different organs. Among other functions, the mouse Six2 is involved in the late differentiation of the retina (Kawakami *et al.*, 1996). The vertebrate Six1/ Six2 homologues do not seem to be involved in the early development of the visual system. This notably differs from what is known from protostome studies. Indeed, protostome *Six1/2* homologues, which are often co-expressed with the *Pax-6* homologue, are known to be important for the early specification of the visual system (Cheyette *et al.*, 1994; Arendt *et al.*, 2002).

Optix is also involved in compound eye development and in forebrain development (Seimiya and Gehring, 2000). In contrast to So, Optix does not present any synergistic interaction with Eya and acts independently of Ey (Seimiya and Gehring, 2000). However, recent evidences suggest that *Optix* may be a direct target of Toy (M.Seimiya, personal communication). Vertebrate Six3/ Six6 are involved in head patterning and in eye formation (Oliver *et al.*, 1995a). In medaka fish, Six3 displays a pivotal role in the development of the retina as a part of a conserved network. Indeed, the *Six3* overexpression has been shown to result in ectopic retina formation (Loosli *et al.*, 1999). In addition, Pax-6 has been shown to directly regulate the expression of *Six3* during the development of the zebrafish visual system (Wargelius *et al.*, 2003). *Six3* is usually expressed in lens and retina structures, while *Six6* is restricted to the retina. In addition to their conserved role during eye development, Six3 and Six6 are also required for the vertebrate forebrain development (Conte *et al.*, 2005).

The other *Drosophila* Six protein, D-Six4, is not involved in the process of eye development but rather plays important roles during the patterning of the head and the development of the nerve cord, the muscles and the gonads (Seo *et al.*, 1999). A role for D-Six4 in the cell recognition events, which are required for myoblast fusion and for the interaction between germline and somatic cells, has been proposed (Kirby *et al.*, 2001). Members of the Six4/Six5 subfamily are commonly involved in the development of sensory structures and in myogenesis. In mouse and human, mutations in *Six4/Six5* genes are associated with myotonic dystrophy (Personius *et al.*, 2005).

1.4.4 The Dach transcription factors

The *Drosophila dac* and the vertebrate homologues, *Dach1* and *Dach2*, encode transcription factors that are characterized by two conserved domains, the DachBox-N and the DachBox-C domains (Fig. 3.8) (Kozmik *et al.*, 1999). Both, in *Drosophila* and vertebrates, Dac proteins are required for the development of the eyes, the brain and the legs (Mardon *et al.*, 1994; Davis *et al.*, 2001). The eyes of *dac* null mutant homozygotes are either absent or severely reduced, where any of the few ommatidia formed have a normal morphology (Mardon *et al.*, 1994).

Dach has been proposed to be a direct target of Pax-6/ ey. Although the DachBox-N and the DachBox-C domains are highly conserved, only the DachBox-N domain seems to be essential for Dac function in *Drosophila* (Tavsanli *et al.*, 2004). The DachBox-N domain has some structural similarities with the winged helix/ forkhead subgroup of the helix-turn-helix DNA-binding protein family (Kim *et al.*, 2002). Even though no specific DNA-binding sites are known for Dach, it binds naked DNA (Ikeda *et al.*, 2002). In addition to its potential role in promoting transcription, a role in transcription repression has also been suggested for Dach (Li *et al.*, 2002). When co-expressed in *Drosophila*, *dac* and *eya* increase both the size and the frequency of ectopic eyes (Chen *et al.*, 1997). Hence, their synergistic activity argues for the idea that Dac and Eya proteins act as a complex during *Drosophila* eye development. Indeed, the DachBox-C domain is known to mediate protein-protein interactions with other RDGN members, like Eya via its ED (Chen *et al.*, 1997).

2. Aim of the second project of this PhD thesis

Sun radiations represent the primary energy source at the basis of life in nearly all ecosystems. It is widely accepted that the light has probably been one of the most important selection pressures during evolution. A tremendous variety of photoreceptive-structures, from simple eyespots to the more elaborated vertebrate camera-type eyes have been generated among living organisms to detect light. The fact that, based on traditional anatomical and embryological studies, the eye-structures of the various animal phyla appear very different, has lead to the idea of an independent evolution of the eyes. However, this view has been greatly challenged by more recent molecular analyses and especially by the discovery that *Pax-6*, a homeobox- and paired box-containing gene, initiates eye formation in all or almost all Bilateria. Furthermore, it has been proposed that the eye-field of various organisms is specified by the coordinated expression of homologous members of an evolutionary conserved network: the so-called retinal determination gene network (RDGN). It is composed of members of the conserved transcription factor families Pax, Six, Eya and Dach, which interact in a complicated manner to specify the bilaterian eyes. Interestingly, it has been shown that not only the composition of the network but also several of the interactions between the network members are conserved between vertebrates and *Drosophila*. This molecular unity, which underlies the development of bilaterian eyes, argues for a monophyletic origin of the eyes.

To find out to which extend the specification of the bilaterian eyes is truly conserved, it is important to investigate Lophotrochozoa, the third clade of bilaterian animals, as most of the molecular data about eye formation stems from Deuterostomia and Ecdysozoa. It is already known that the expression pattern of *Pax-6* correlates with the location of the newly forming eyes, during *Lineus* development and regeneration. Additionally, *Pax-6* remains expressed in the adult worms. We have shown recently that its inactivation by RNAi leads to the disappearance of the differentiated eyes. This suggests that *Pax-6* is probably not only involved in the development and the regeneration of the *Lineus* eyes, but also in their adult maintenance.

The aim of this second project of this thesis work was to further investigate the specification of the *Lineus* visual structures. We wanted to know whether homologues of the RDGN members are involved in the development, maintenance and regeneration of the pigment cup- eyes of *Lineus.*

We took the following approaches to answer these questions:

- Based on functional analysis and expression patterns in vertebrates and *Drosophila*, it is conceivable that the *Six* genes may be involved in the lophotrochozoan eye specification. Are there homologues of the *Six* gene family in the *Lineus* genome? If so, what kind of expression patterns do the *LsSix* homologues have? Is there one, among them, which would suggest an involvement of a *LsSix* gene in the *Lineus* eye specification during development and regeneration?

- In order to investigate the development, maintenance and regeneration of the *Lineus* eyes, we need to find a gene that is constantly expressed in these structures. With such a gene at hand, we could follow the fate of the *Lineus* eyes, after gene activation by RNAi. The *Opsin* genes are good candidates for such a purpose. Therefore, is there any *opsin* gene in the *Lineus* genome? If so, is it constantly expressed in the eyes, as expected for an *opsin* homologue?

3. Results

3.1 *LsPax* genes

3.1.1 Characterization of the full-length *LsPax-6* and identification of a *LsPax-6* splice variant

A *LsPax-6* homologue, which shares extensive sequence identity and several conserved splice sites with the mammalian and *Drosophila* genes**,** has already been isolated from *L.sanguineus* (Loosli *et al.*, 1996). During regeneration of *L.sanguineus* heads, *LsPax-6* is expressed in the CNS, in the cerebral organs and in the eye region (Loosli *et al.*, 1996). In addition, it has been shown to be expressed in fully regenerated eyes and in intact eye regions (Tarpin *et al.*, 1999). Furthermore, the transposition of *Lineus* postocellar tissue cells into ocellar location is known to cause unexpected *LsPax-6* expression, which results in eye development (Tarpin *et al.*, 2002). On top, we have previously reported that inactivation of the *LsPax-6* by RNA-mediated gene interference (RNAi) in an adult *L.sanguineus* leads to the disappearance of the eyes, and blocks the regeneration of the head from an amputated worm (Charpignon, 2002). These RNAi induced phenotypes are transient. When the *LsPax-6* dsRNA injection was ceased, the eyes reappeared within few days and the regeneration process began in head-amputated *L.sanguineus* (Charpignon, 2002).

To further investigate the biochemical activities and gene regulatory functions of *Ls-Pax6*, we plan to test, whether ectopic expression of *Ls-Pax6* in *Drosophila* imaginal discs can induce supernumerary eye structures, as it has been observed for invertebrate and vertebrate *Pax-6* homologues. When comparing the available LsPax-6 protein sequence to several other Pax-6 homologues, we noticed that a highly conserved proline, serine and threonine-rich (PST) domain at the C-terminus was apparently not present in the LsPax-6 protein sequence. Therefore, to make sure we would inject a full-length *LsPax-6* sequence into *Drosophila* and would not miss its 3'end, we decided to perform a 3' RACE PCR on *L.sanguineus* cDNA made from mRNA extracted from worms undergoing head regeneration. We used a specific *Ls-Pax6* primer, which corresponds to a sequence located in the homeobox. It appears that the published *LsPax-6* sequence (see appendix 2.) contains mistakes and is incomplete at its C-terminus. This is probably due to a frame shift during the sequencing analysis of the gene, leading to a mis-positionning of an exon/ intron boundary.

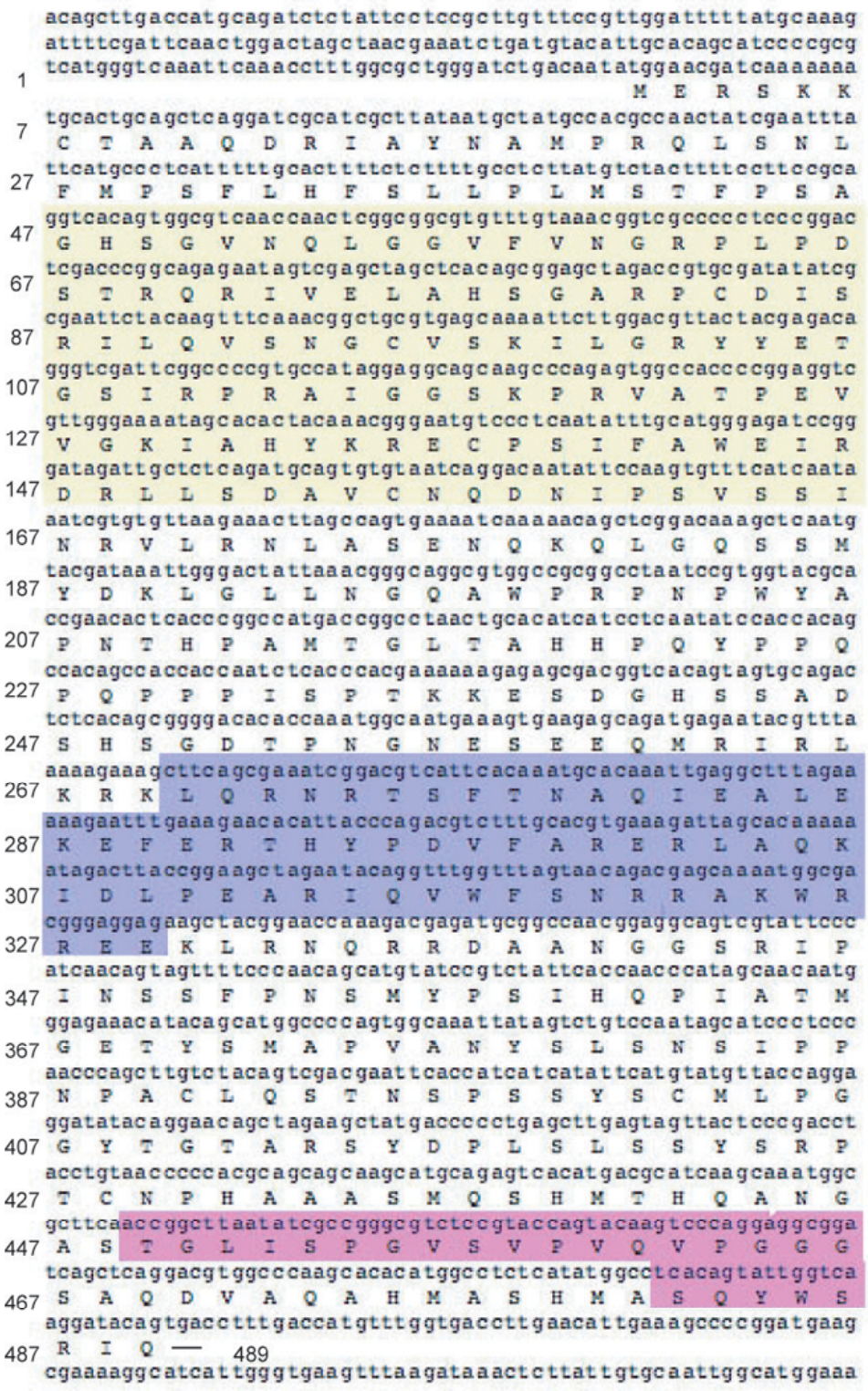


Fig. 3.10 The full-length *LsPax-6* sequence and the deduced amino acid sequence of its putative ORF

The putative *LsPax-6* encodes a 489 amino acids protein. The PD and its corresponding nucleotidic sequence are highlighted in yellow, while the HD and its corresponding nucleotidic sequence are highlighted in blue. Two highly conserved regions of the newly identified PST domain are highlighted in pink.

Sequencing of the *LsPax-6* RACE PCR product revealed the presence of a previously un-identified C-term motif in the deduced *LsPax-6* protein sequence (Fig. 3.10). Hence, the correct LsPax-6 protein sequence is longer than the one published (489 aa instead of 370 aa) and contains the highly conserved PST domain at its C-terminus.

In addition, when sequencing the 3' RACE PCR product of *LsPax-6*, we found a variant mRNA sequence (Fig. 3.12). To make sure of the existence of this variant *LsPax-6* mRNA, we repeated the 3' RACE PCR on three distinct *L.sanguineus* cDNA preparations from several animals. We could find it in all three distinct cDNA preparations. Then, we made PCR on *L.sanguineus* genomic DNA and found that this other *LsPax-6* mRNA corresponds to an alternative splice variant. When comparing both mRNA sequences, we found that the alternative splice variant possesses an "insertion" of 14 nucleotides. This is due to the presence of two possible 5' donor sites in the *Ls-Pax6* exon 4 (Fig. 3.11). These two donor sites are separated by exactly 14 nucleotides (see appendix 3.). When the second splicing donor site (D') is used, the resulting mRNA is 14 nucleotides longer compared to the one resulting from splicing using the first splicing donor site (D). Those 14 additional nucleotides, present in the *LsPax-6* splice variant, are responsible for a frame shift, which leads to the creation of an earlier stop codon in the LsPax-6 protein. Thus, the hypothetical protein deduced from the alternative *LsPax-6* splice is shorter (455 aa instead of 489) and truncated: interestingly, it lacks the majority of the highly conserved region of the PST domain at its C-terminus (See also appendix 4.).

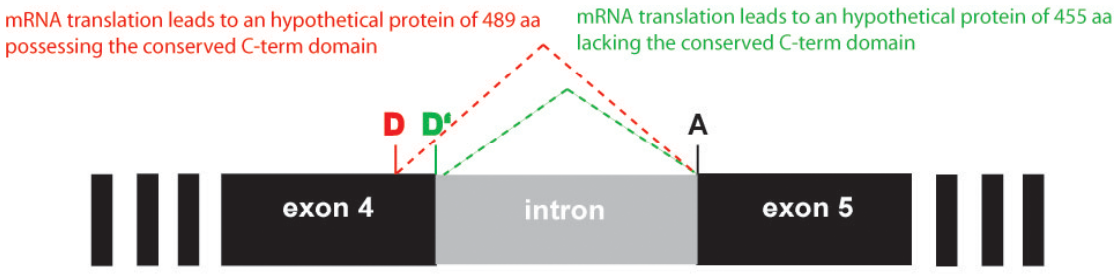


Fig. 3.11 Scheme representing the *LsPax-6* alternative splicing process
D: 5' donor site; D': alternative 5' donor site; A: 3' acceptor site. D is separated from D' by only 14 nucleotides.
From genomic DNA sequencing, it appears that two donor sites are present in exon 1. When the 3' donor site D is used for the splicing, the deduced protein sequence from the mRNA is composed of 489 aa and possesses the highly conserved C-term domain. When the 3' donor site D' is used for the splicing, the deduced protein sequence from the mRNA is composed of 455 aa and lacks completely the highly conserved C-term domain.

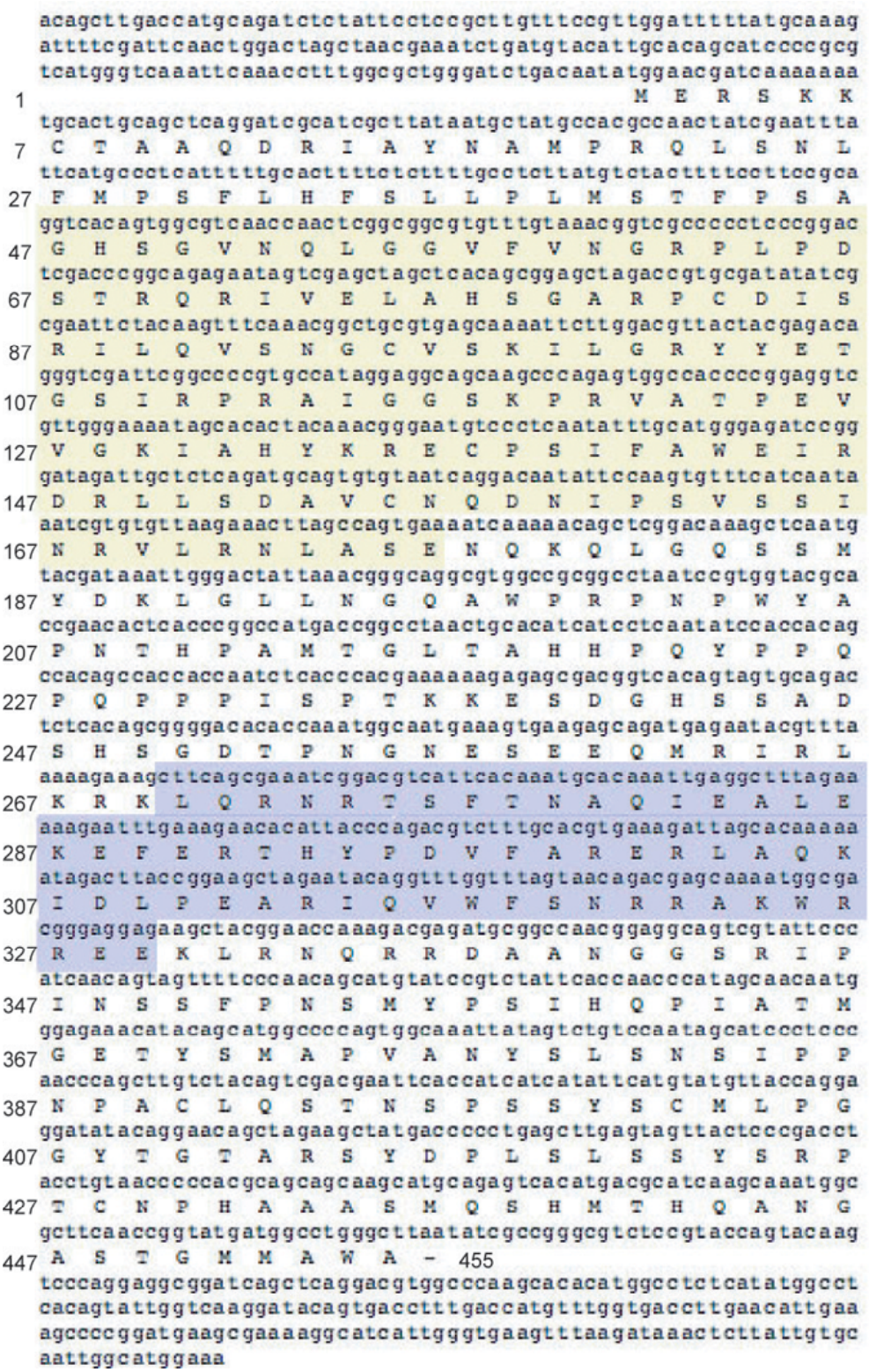


Fig. 3.12 The deduced amino acid sequence of the putative ORF of the alternative splice

The alternative *LsPax-6* splice encodes a truncated Ls-Pax6 protein of 455 amino acids protein. The PD and its corresponding nucleotidic sequence are highlighted in yellow, while the HD and its corresponding nucleotidic sequence are highlighted in blue. Notice the absence of the evolutionary conserved PST domain at the C-terminus of the protein.

Also the vertebrate *Pax-6* locus is known to encode two products caused by alternative splicing of exon 5a that adds an additional 14 aa residues within the paired domain (Epstein *et al.*, 1994). As the insertion of these 14 additional aa disrupts the PAI subdomain, the longer PD of the 5a isoform preferentially interacts with DNA through its RED subdomain. Thus, the Pax6(5a) protein exhibits unique DNA-binding properties (Kozmik *et al.*, 1997). *Pax6(5a)* is expressed in the vertebrate eye at about one-tenth the level of *Pax6* during embryonic development (Kozmik *et al.*, 1997). The ratio of these two Pax6 isoforms has been shown to be critical for the normal development and function of vertebrate eyes (Singh *et al.*, 2002). Thus, alternative splicing of *Pax-6* locus might be important for Pax-6 function. The presence of evolutionary conserved splice sites within the *LsPax-6* transcript has been reported (Loosli *et al.*, 1996). Interestingly, one splice site is found in the paired box, at the same position as in the human, mouse, quail, *Drosophila*, and *C. elegans* sequences. In order to test the possible existence of an alternative splicing mechanism, which results in truncation of the PD, we performed RT-PCR experiments, using specific primers flanking the paired box, on mRNA extracted from various adult and regenerating tissues. The RT-PCR products were gel-analyzed and sequenced. However, no other additional *LsPax-6* splice variant was identified.

3.1.2. Identification of a *LsPax-2/5/8* gene

In our quest for *LsPax-6* splice variants, we have unexpectedly identified a gene fragment, which encodes a Pax-2/5/8-type PD. When sequencing the isolated *Pax-2/5/8*-type paired box, we realized that our *LsPax-6* primers, specific for the paired box, have mismatched with the paired box of a putative *LsPax-2/5/8* gene. The *Pax-2/5/8*-type paired box fragment was subsequently extended by 3' and 5' RACE PCR procedures using specific primers. Sequencing of the RACE PCR fragments allowed us to identify this gene as a *LsPax-2/5/8* gene. We isolated a full length *LsPax-2/5/8*, which encodes a protein of 401 aa. It contains a Pax-2/5/8-type PD, a conserved octapeptide and some residues of a HD (Fig. 3.13). The PD of LsPax-2/5/8 exhibits three amino acids characteristics of Pax-2/5/8-type PD: Q42, R44 and H47 (Fig. 3.14). These amino acids are known to be responsible for differences in the DNA-binding capacities between Pax-2/5/8 proteins and Pax-6 proteins.

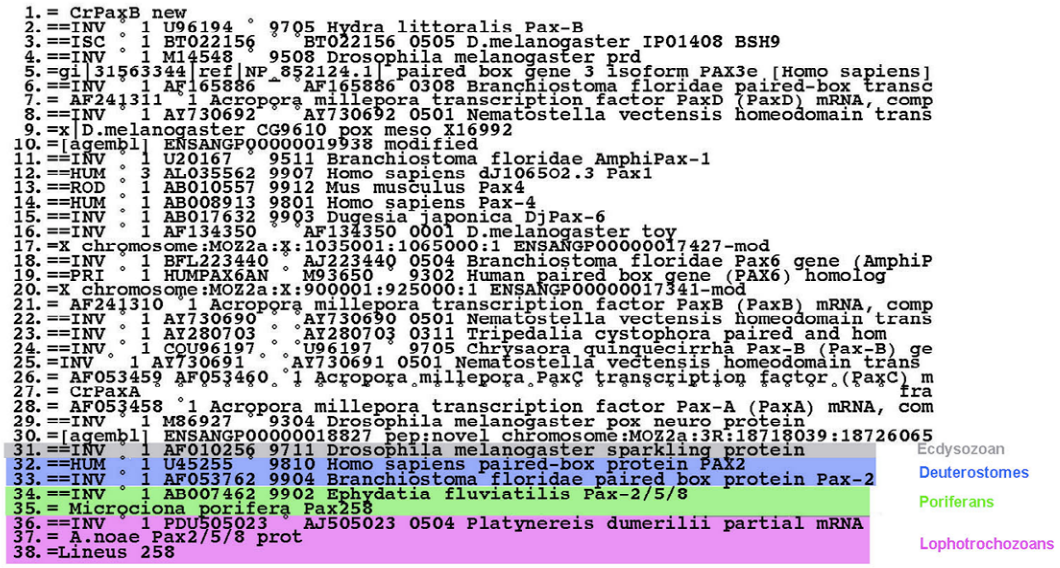
Fig. 3.13 The full-length *LsPax-2/5/8* sequence and the deduced amino acid sequence of its putative ORF

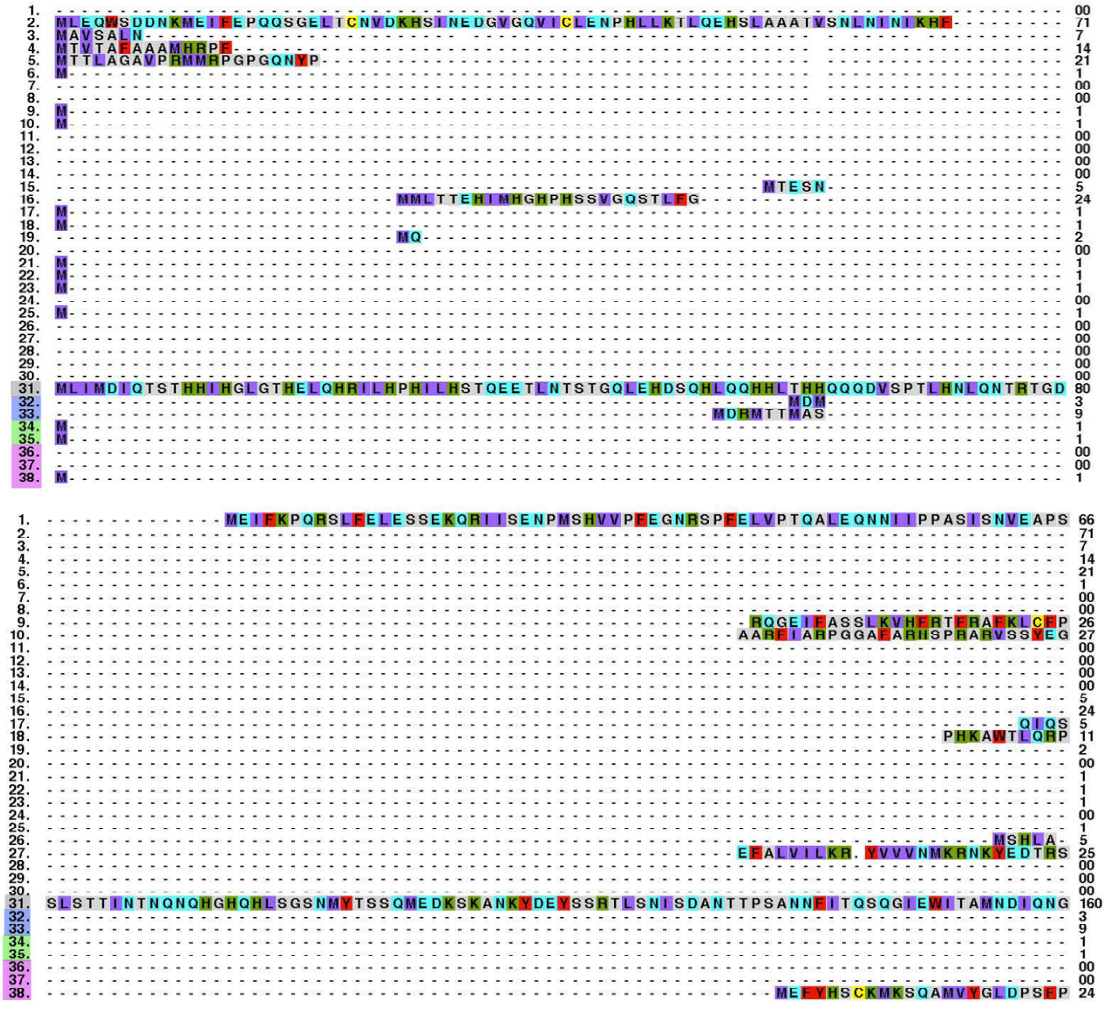
The putative *LsPax-2/5/8* encodes a 401 amino acids protein. The PD and its corresponding nucleotidic sequence are highlighted in yellow; the octapeptide and its corresponding nucleotidic sequence are highlighted in blue.

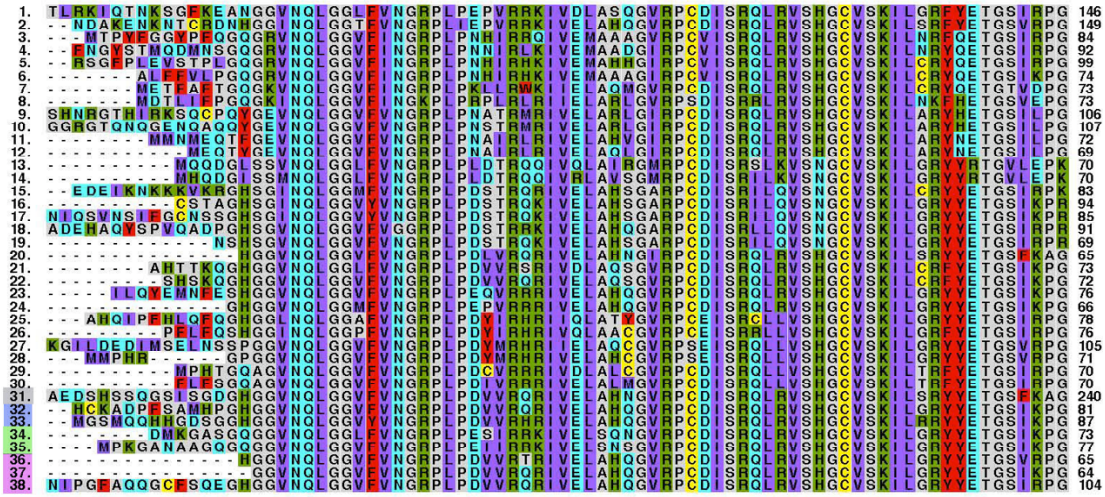
```
LsPax-2/5/8  VNQLGGVFVNGRPLPDVVRQRIVELAHQGVRPCDISRQLRVSHGCVSKILGRYYETGSIR
             VNQLGGVFVNGRPLPD  RQRIVELAH G RPCDISR L+VS+GCV+KILGRYYETGSIR
LsPax-6      VNQLGGVFVNGRPLPDSTRQRIVELAHSGARPCDISRILQVSNGCVTKILGRYYETGSIR

LsPax-2/5/8  PGVIGGSKPKVAXPKVVDAILHYKAESPTMFAWEIRDMLLSECVCSQENVPSVXSINRIV
             P  IGGSKP+VA P+VV  I HYK E P++FAWEIRD LLS+ VC+Q+N+PSV SINR++
LsPax-6      PRAIGGSKPRVATPEVVGKIAHYKRECPSIFAWEIRDRLLSDAVCNQDNIPSVSSINRVL

LsPax-2/5/8  RN
             RN
LsPax-6      RN
```

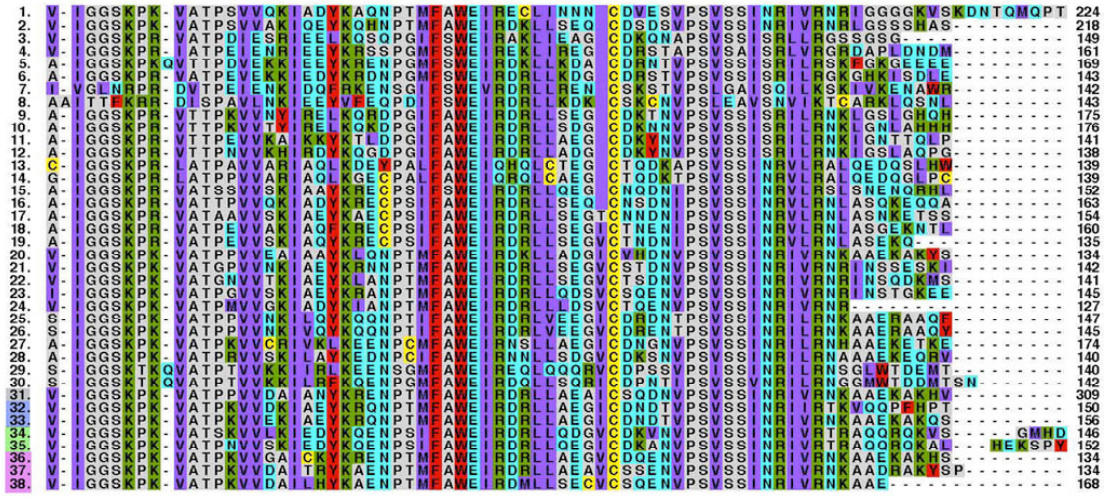
Fig. 3.14 Comparison of the LsPax-2/5/8 PD with the LsPax-6 PD

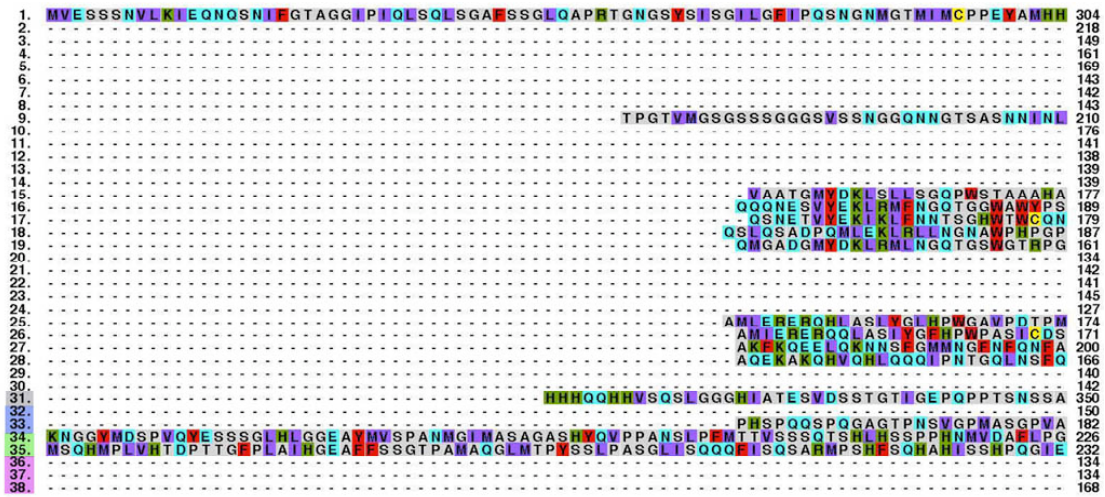
Three amino acids, within the PAI subdomain of the PD, are used to distinguish Pax-2/5/8-type PD from Pax-6-type PD. These aa are located at position 42, 44 and 47 of the PD and are responsible for different DNA-binding specificities. The PD of LsPax-2/5/8 exhibits a Q42, a R44 and a H47, whereas the PD of LsPax-6 exhibits an I42, a Q44 and a N47.

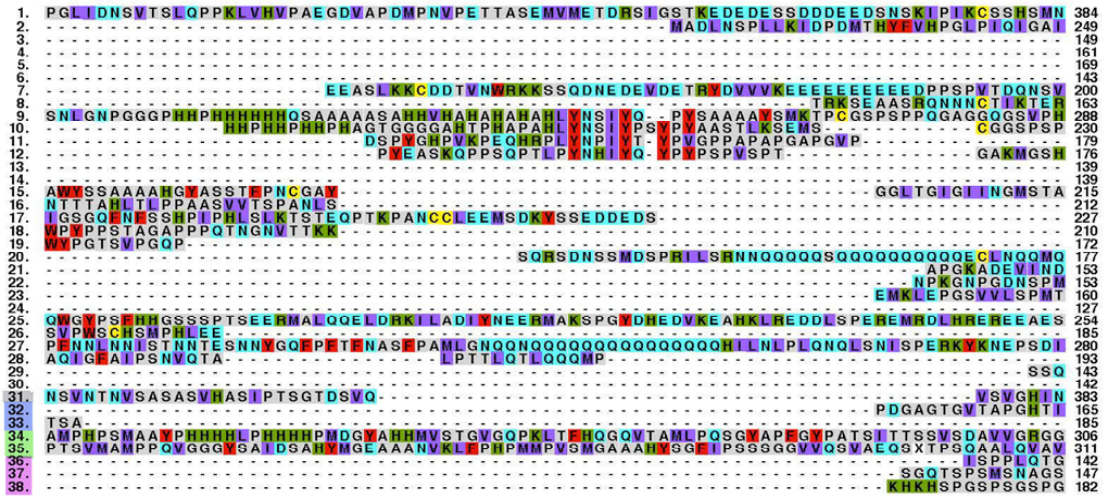

1. = CrPaxB new
2. ==INV ° 1 U96194 ° 9705 Hydra littoralis Pax-B
3. ==ISC ° 1 BT022156 ° BT022156 0505 D.melanogaster IP01408 BSH9
4. ==INV ° 1 M14548 9508 Drosophila melanogaster prd
5. =gi|31563344|ref|NP_852124.1| paired box gene 3 isoform PAX3e [Homo sapiens]
6. ==INV ° 1 AF165886 ° AF165886 0308 Branchiostoma floridae paired-box transc
7. = AF241311 °1 Acropora millepora transcription factor PaxD (PaxD) mRNA, comp
8. ==INV ° 1 AY730692 ° AY730692 0501 Nematostella vectensis homeodomain trans
9. =x|D.melanogaster CG9610 pox meso X16992
10. =[agembl] ENSANGP00000019938 modified
11. ==INV ° 1 U20167 ° 9511 Branchiostoma floridae AmphiPax-1
12. ==HUM ° 3 AL035562 9907 Homo sapiens dJ106502.3 Pax1
13. ==ROD ° 1 AB010557 9912 Mus musculus Pax4
14. ==HUM ° 1 AB008913 9801 Homo sapiens Pax-4
15. ==INV ° 1 AB017632 9903 Dugesia japonica DjPax-6
16. ==INV ° 1 AF134350 ° AF134350 0001 D.melanogaster toy
17. =X chromosome:MOZ2a:X:1035001:1065000:1 ENSANGP00000017427-mod
18. ==INV ° 1 BFL223440 ° AJ223440 0504 Branchiostoma floridae Pax6 gene (AmphiP
19. ==PRI ° 1 HUMPAX6AN ° M93650 ° 9302 Human paired box gene (PAX6) homolog
20. =X chromosome:MOZ2a:X:900001:925000:1 ENSANGP00000017341-mod
21. = AF241310 °1 Acropora millepora transcription factor PaxB (PaxB) mRNA, comp
22. ==INV ° 1 AY730690 ° AY730690 0501 Nematostella vectensis homeodomain trans
23. ==INV ° 1 AY280703 ° AY280703 0311 Tripedalia cystophora paired and hom
24. ==INV ° 1 CQU96197 ° U96197 ° 9705 Chrysaora quinquecirrha Pax-B (Pax-B) ge
25. =INV ° 1 AY730691 ° AY730691 0501 Nematostella vectensis homeodomain trans
26. = AF053459 AF053460 °1 Acropora millepora PaxC transcription factor (PaxC) m
27. = CrPaxA fra
28. = AF053458 °1 Acropora millepora transcription factor Pax-A (PaxA) mRNA, com
29. ==INV ° 1 M86927 9304 Drosophila melanogaster pox neuro protein
30. =[agembl] ENSANGP00000018827 pep:novel chromosome:MOZ2a:3R:18718039:18726065
31. ==INV ° 1 AF010256 9711 Drosophila melanogaster sparkling protein
32. ==HUM ° 1 U45255 ° 9810 Homo sapiens paired-box protein PAX2
33. ==INV ° 1 AF053762 9904 Branchiostoma floridae paired box protein Pax-2
34. ==INV ° 1 AB007462 9902 Ephydatia fluviatilis Pax-2/5/8
35. = Microciona porifera Pax258
36. ==INV ° 1 PDU505023 ° AJ505023 0504 Platynereis dumerilii partial mRNA
37. = A.noae Pax2/5/8 prot
38. =Lineus 258
Ecdysozoan
Deuterostomes
Poriferans
Lophotrochozoans

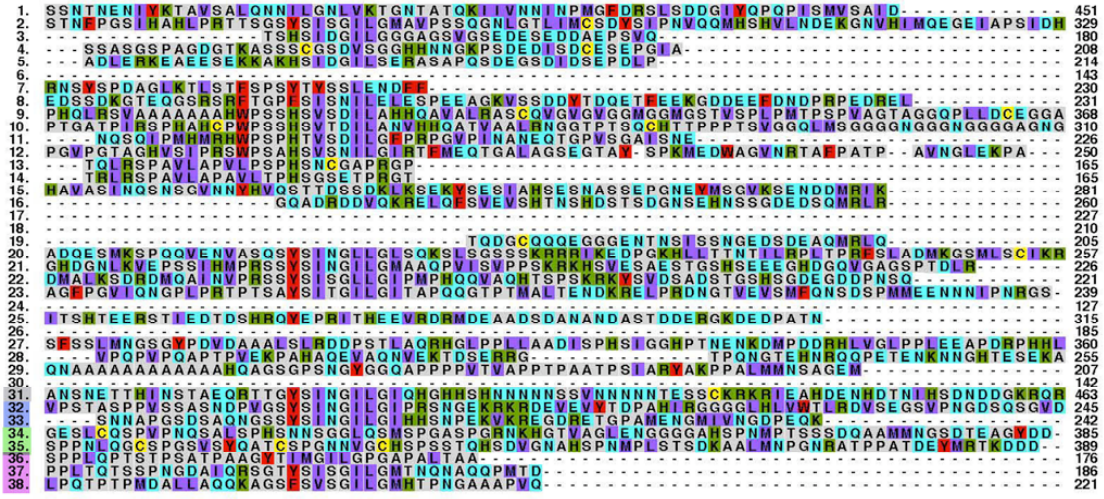


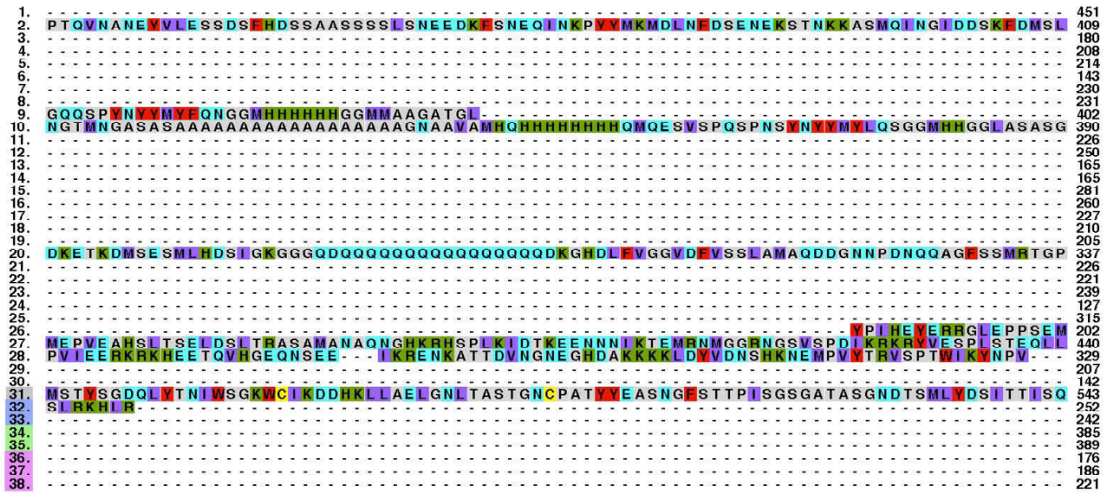


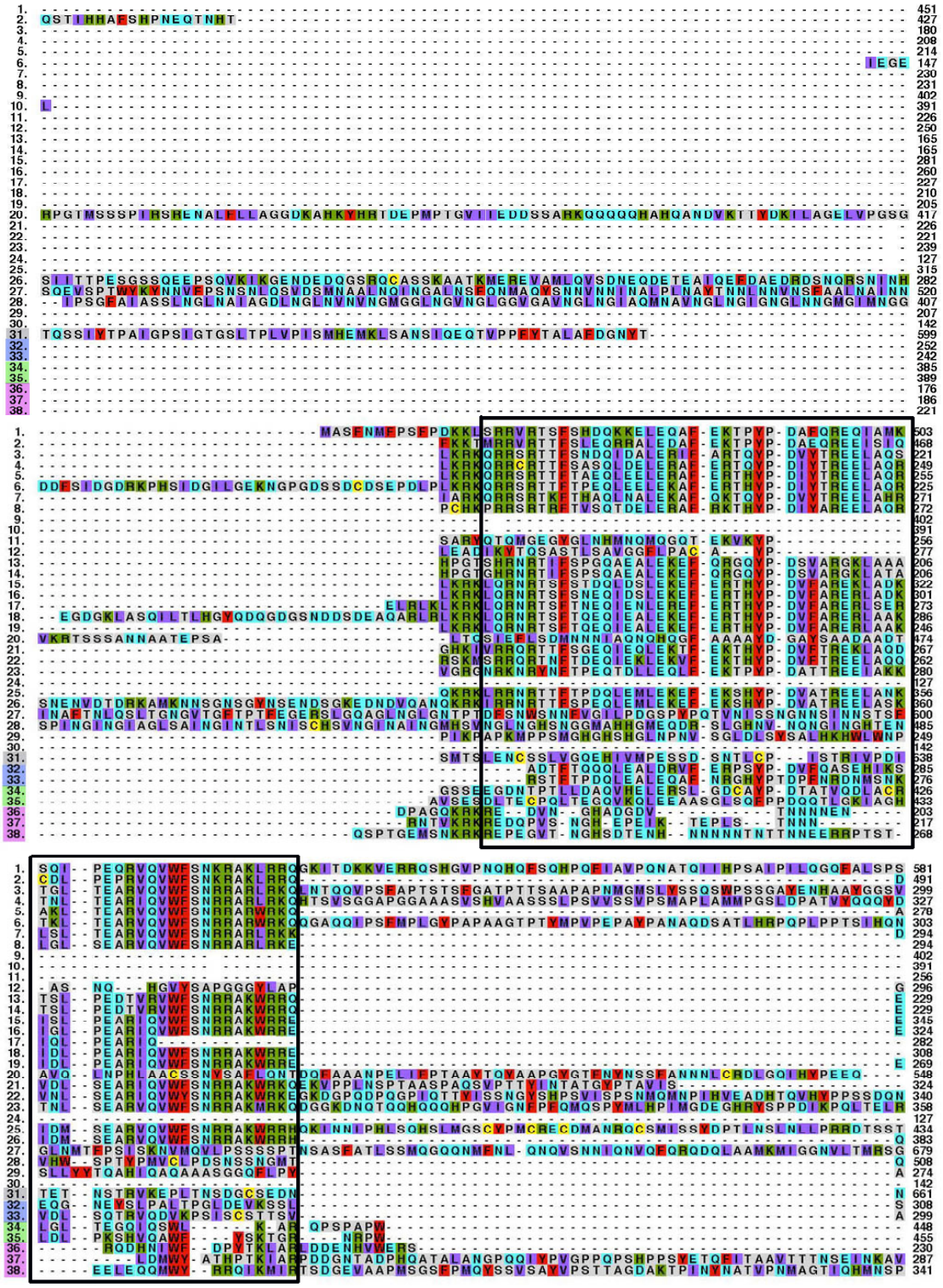












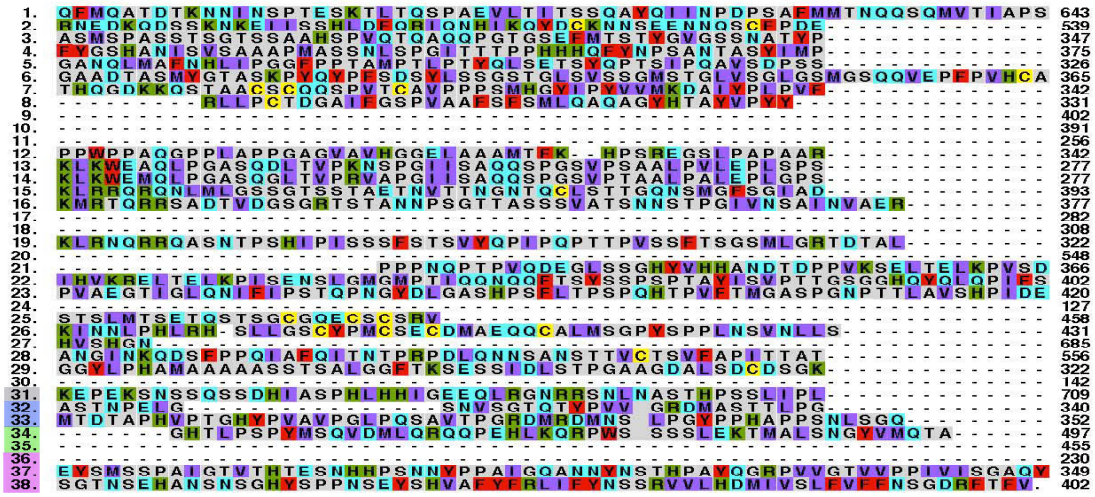


Fig. 3.15 Alignment of LsPax-2/5/8 protein with other Pax family members
We have aligned the LsPax-2/5/8 protein sequence (sequence number "38") with other Pax-2/5/8 homologues and members of other Pax subfamilies. Pax-2/5/8 protein sequences from Lophotrochozoa are highlighted in pink: "36" corresponds to the polychaete *Platynereis dumerili* sequence, "37" corresponds to the mollusk *Arca noae* sequence and "38" corresponds to the *Lineus sanguineus* sequence. A Pax-2/5/8 protein from an ecdysozoan member is highlighted in grey and corresponds to the Sparkling protein sequence of *Drosophila melanogaster*, number "31". Pax-2/5/8 protein sequences from Deuterostomia are highlighted in blue: "32" corresponds to the *Homo sapiens* Pax-2 sequence and "33" corresponds to the *Branchiostoma floridae* Pax-2 sequence. In addition to Pax-2/5/8 protein sequences from bilaterian animals, we have added homologues of two species of sponges (poriferans). They are highlighted in green: "34" corresponds to the *Ephydatia fluviatilis* sequence and "35" corresponds to the *Microciona porifera* sequence.
The region corresponding to the alignment of the HDs is boxed. Note that deuterostome and ecdysozoan Pax-2/5/8-type HDs are partial (sequences "32" and "33") and contain only residues from the first helix. Lophotrochozoan Pax-2/5/8-type HDs are also partial, but contain conserved residues from the third helix (indicated by pink asterisks). In contrast, although divergent, the sponge Pax-2/5/8-type HDs seem to be complete.

The alignment of the LsPax-2/5/8 protein sequence with other sequences of Pax-2/5/8 homologues revealed an interesting feature of the lophotrochozoan homologues. From identification of various Pax-2/5/8-related proteins from bilaterian animals, the Pax-2/5/8-type HD is known to be a partial one. So far, it was thought to be composed only of residues from the first helix. However, when comparing Pax-2/5/8-type HDs from lophotrochozoan animals, it seems that the first helix of the HD is not conserved at all, whereas residues from the third helix are present (Fig. 3.15). In respect with sequences available from databases, like Genbank, only three Pax-2/5/8-type HD have been identified from lophotrochozoan animals: one from the polychaete *Platynereis*, one from the mollusk *Arca* (L.Keller, personal communication) and one identified during this PhD thesis work, from the nemertean *Lineus*.

3.2. Isolation of a *Ls-opsin* gene and characterization of its expression pattern

As already mentioned above, we have previously reported that inactivation of the *LsPax-6* by RNA-mediated gene interference (RNAi) in an adult *L.sanguineus* leads to the disappearance of the eyes (Charpignon, 2002). However, this assumption was based only on the fact that we were not able to detect anymore the presence of the usual brown pigment, which is used to determine the location of the *Lineus* eyes. Hence, we decided to search for a molecular marker for *Lineus* eyes. An *opsin* homologue appeared to be a marker of choice since all animal eyes contain rhodopsin.

Similarly to other G protein-coupled receptors (GPCRs), opsin proteins possess seven transmembrane helice. But they can be distinguished from other G-protein coupled receptors (GPCRs) by the specific presence of a lysine residue (K) in their 7^{th} transmembrane helix. This K binds retinal via a Schiff-base linkage: the nitrogen atom of the K forms a double bond with the carbon atom at one end of the retinal. To clone an *opsin* homologue from *L.sanguineus*, we used a first set of degenerated PCR primers designed against the sequence encoding the 7^{th} transmembrane helix. The PCR product, resulting from this first round of PCR, was subsequently used for hemi-nested PCR. We used a degenerated primer against the region coding for the seventh helix, which comprises the K specific for opsin-type GPCRs. This procedure allows us to discriminate opsin-type GPCRs from the various members of the GPCR superfamily. Sequencing of the subsequent RACE PCR products revealed that we had identified a full-length *ospin* homologue from *L.sanguineus* (Fig. 3.16). It encodes a protein of 332 aa, which, similarly to other GPCRs, exhibits seven transmembrane helices. The amino acids sequence of Ls-opsin contains a K in its seventh helix (Fig. 3.16) confirming that this newly found G protein-coupled receptors is truly an opsin. From phylogenetic studies, it appears that Ls-opsin sequence does not cluster with classical rhabdomeric-type opsin and probably corresponds to a new type of invertebrate opsin (see appendix 5.).

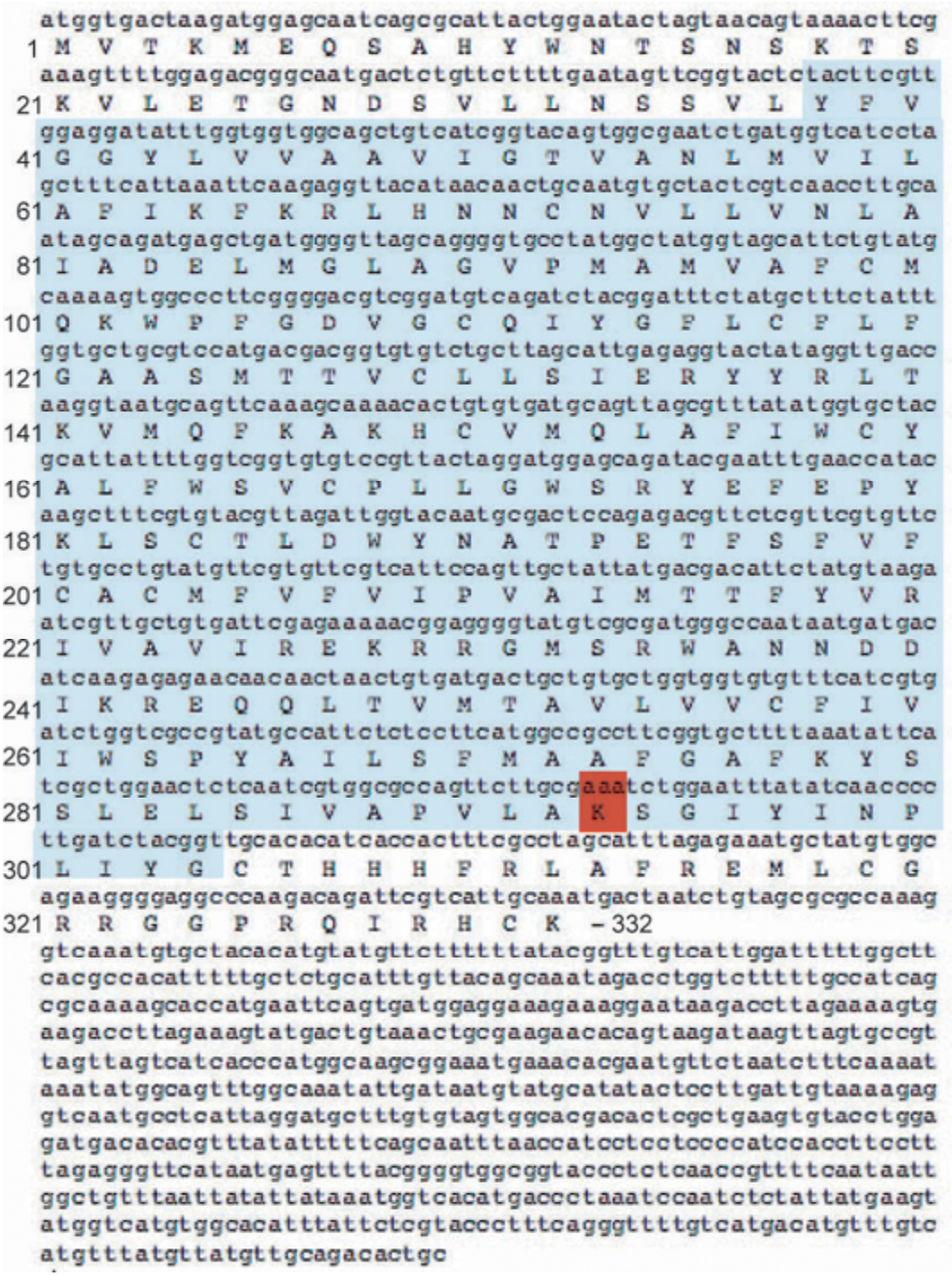


Fig. 3.16 The full-length *Ls-opsin* sequence and the deduced amino acid sequence of its putative ORF

The *Ls-opsin* encodes a 332 amino acids protein. The seven-transmembranes domain is highlighted in blue. It contains a lysine (K), which is the retinal-binding site, in its seventh transmembrane domain.

We investigated opsin expression during *Lineus* development, maintenance of adult eyes and eye regeneration. Using a *Ls-opsin* probe, we could observe a strictly restricted expression pattern at the level of the developing eyes of *L.viridis* (Fig. 3.17). But we didn't observe expression of *Ls-opsin* neither in the adult *L.sanguineus* eyes nor in regenerating *L.sanguineus* eyes (not shown).

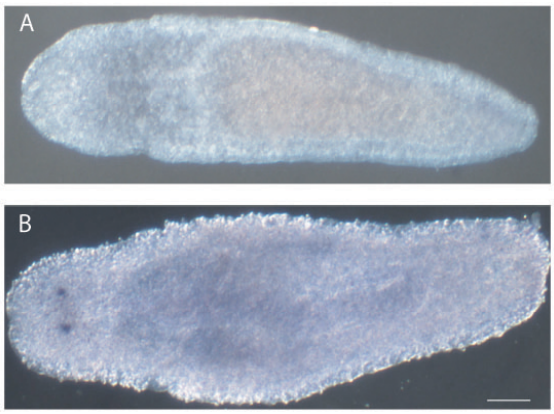


Fig. 3.17 Expression analysis of *Lv-opsin* in a developing *L.viridis* larva
Whole-mount ISH were done using *Ls-opsin* sense (A) and anti-sense (B) RNA probes. Anterior is oriented to the left, dorsal view. The larvae are two-weeks old. Scale bar: 200 µm.

3.3 the *Ls-Six* genes

3.3.1 Cloning and characterization of the *Ls-Six* genes

We used a degenerated primer-based PCR approach in our quest for *Six* genes from *L.sanguineus* (*Ls-Six*). Similarly to the strategy used for cloning of *Ls-Otx*, we aligned the amino acid sequences of different Six homologues in order to deduce evolutionarily conserved regions. We found an almost invariable motif of five amino acids, comprising the three last aa of the SD and the two first aa of the Six-type HD: IWDGE. Hence, we decided to perform a first PCR with the primers Six-GeneralForward, which corresponds to the sequence TIWDGE and Six-HDReverse, which corresponds to the highly conserved region NWFKNRRQ, present in the C-terminal part of Six-type HDs. From this first PCR round, we isolated a 174 bp fragment of a *Six1/2*-type homeobox. A tetrapeptide motif, found in the helix I of the Six-type HD, is commonly used to distinguish between the three subfamilies of Six proteins (Fig. 3.18). As expected, our Six1/2-type HD exhibits an "ETSY motif", while Six4/5-type HD exhibits an "ETVY motif". A less similar motif is found in Six3/6-type HDs, which possess a "QKTH motif".

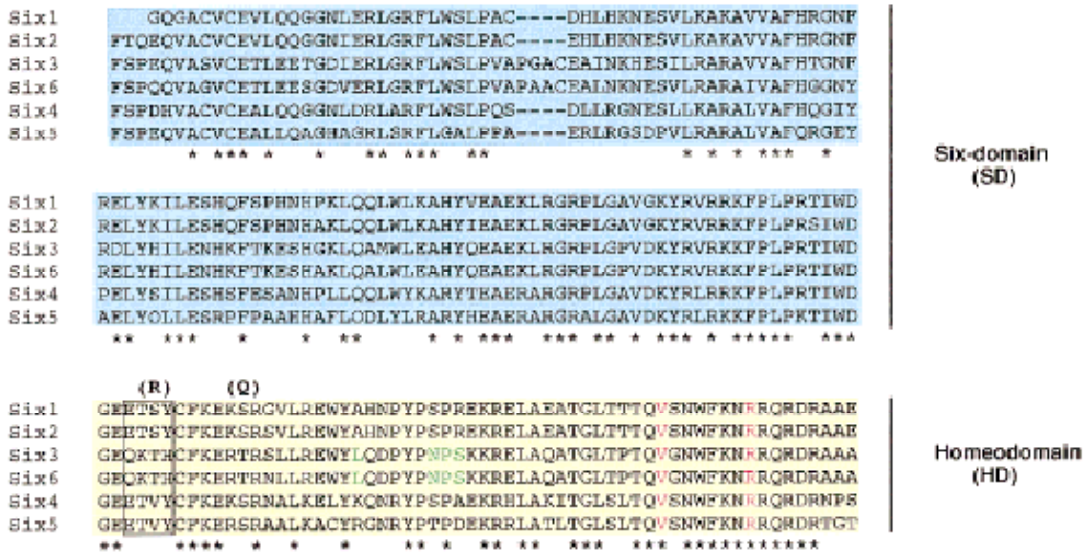


Fig. 3.18 Sequence alignments of the SDs and HDs of mouse Six proteins
The tetrapeptides, identified as "diagnostic residues", are boxed. Two highly conserved residues (R5 and Q12 in helix I) typical of most HDs are shown above the Six-type HDs. Gaps, introduced to maximize sequence homology, are represented by dashes. Asterisks indicate identical amino acids.

Hence, we took advantage of the presence of such "diagnostic residues" to design corresponding degenerated primers. We performed hemi-nested PCR on the PCR product of the first PCR round with two different primer combinations: either the Six3/6Forward (corresponding to the sequence WDGEQKT) and the Six-HDReverse primers or the Six4/5Forward (sequence WDGEET) and the Six-HDReverse primers were used. The first combination of primers leads to the isolation of a 174 pb fragment of a *Six3/6*-type homeobox, while a 174 pb fragment of a *Six4/5*-type homeobox was obtained by the second combination of primers. The three different *Six*-type homeobox fragments were subsequently extended by 3' and 5' RACE procedures, using specific primers. Sequencing the resulting RACE PCR products allows us to identify three different Six genes from *L.sanguineus*. The full-length sequence has been obtained for every one of these genes. Based on sequence comparisons, we were able to name these genes: *Ls-Six1/2*, *Ls-Six3/6* and *Ls-Six4/5* (Fig. 3.20, 3.21, 3.22). Phylogenetic analysis has confirmed their classification into the three main subfamilies (Fig. 3.19).

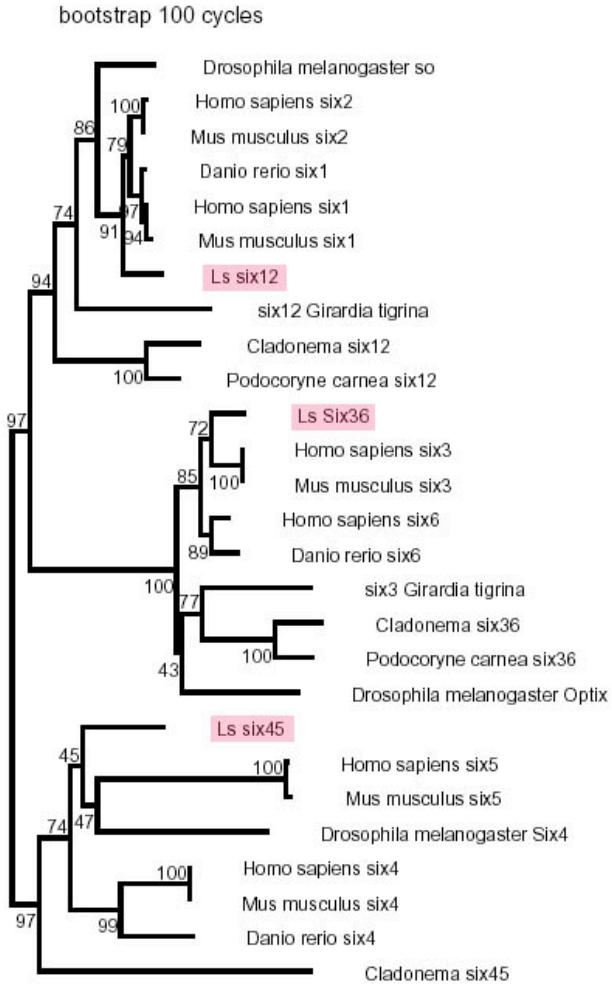


Fig. 3.19 Phylogenetic tree showing *Ls-six* genes in relation to other members of the *Six* gene family

The phylogenetic tree was constructed by using the neighbour-joining (NJ) method. All the amino acids from the full SD and HD were used. Poisson-correction distance was used to draw the NJ tree. The *Ls-Six* genes are highlighted in pink.

Fig. 3.20 The full-length *Ls-Six1/2* sequence and the deduced amino acid sequence of its putative ORF

The putative *Ls-Six1/2* encodes a 294 amino acids protein. The SD and its corresponding nucleotidic sequence are highlighted in violet, while the Six1/2-type HD and its corresponding nucleotidic sequence are highlighted in blue.

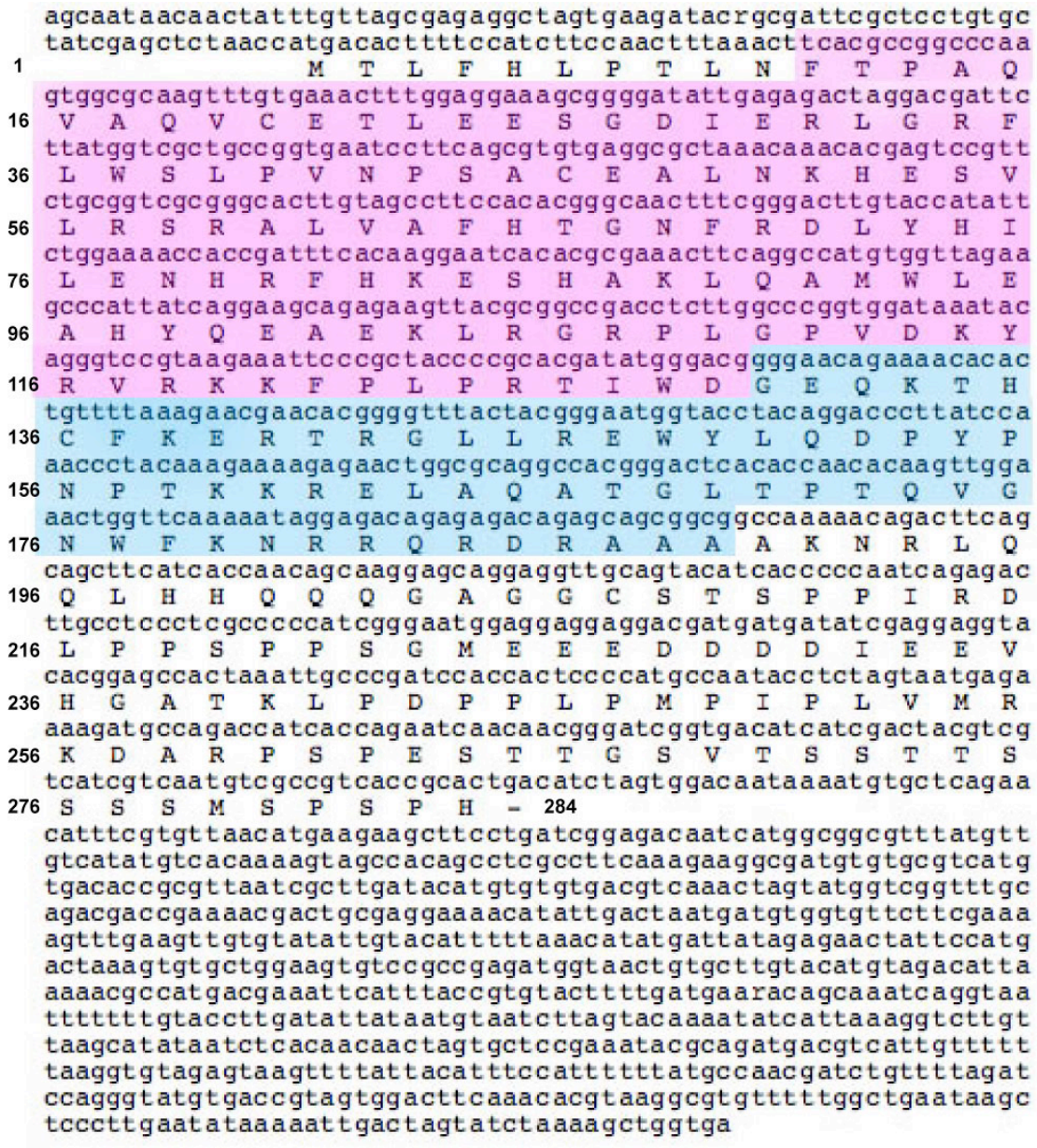


Fig. 3.21 The full-length *Ls-Six3/6* sequence and the deduced amino acid sequence of its putative ORF

The putative *Ls-Six3/6* encodes a 284 amino acids protein. The SD and its corresponding nucleotidic sequence are highlighted in violet, while the Six3/6-type HD and its corresponding nucleotidic sequence are highlighted in blue.

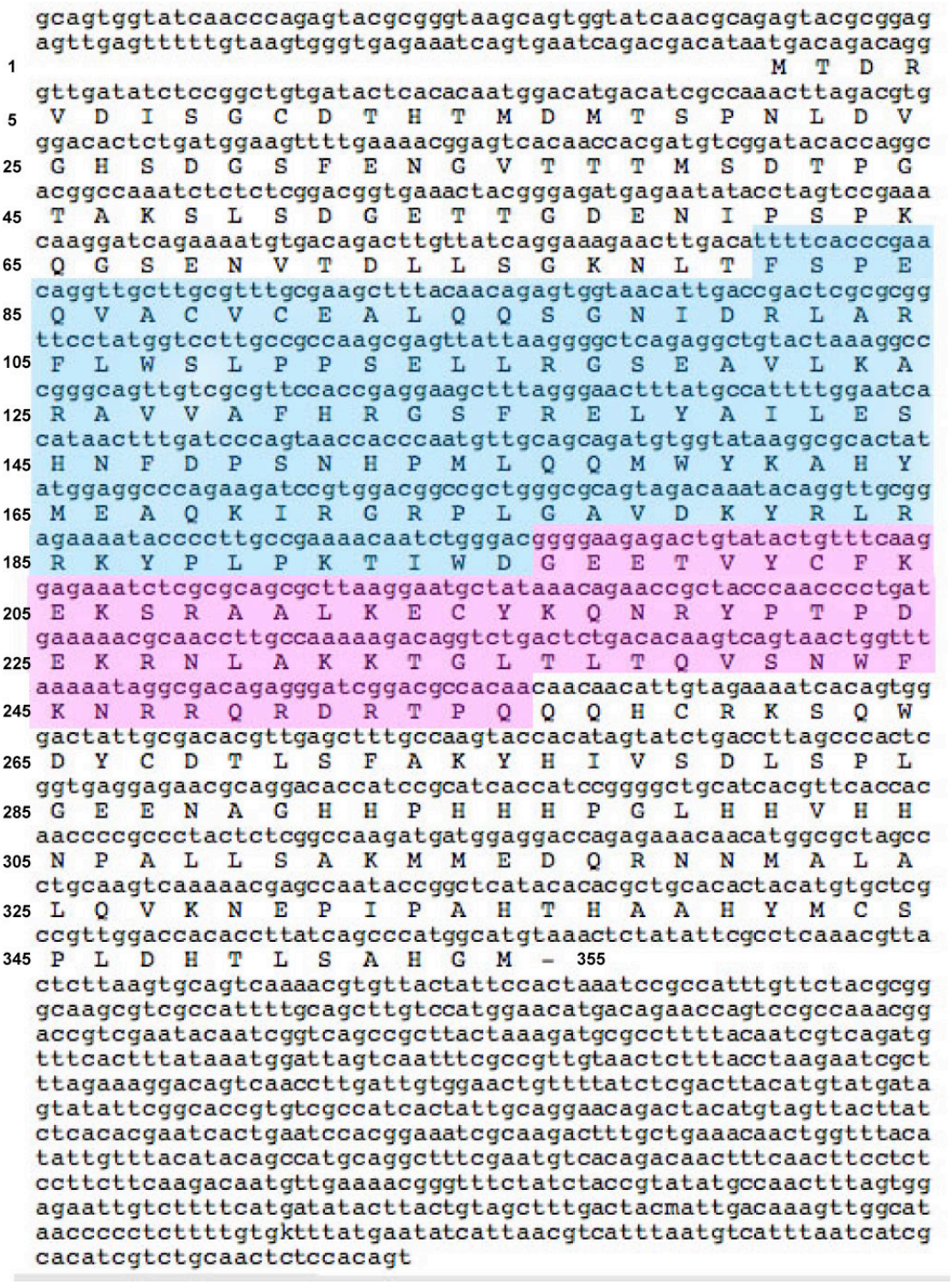


Fig. 3.22 The full-length *Ls-Six4/5* sequence and the deduced amino acid sequence of its putative ORF

The putative *Ls-Six4/5* encodes a 355 amino acids protein. The SD and its corresponding nucleotidic sequence are highlighted in violet, while the Six4/5-type HD and its corresponding nucleotidic sequence are highlighted in blue.

Expression pattern of the *Ls-Six* genes

3.3.2 *Ls-Six1/2* expression patterns

Using a similar strategy than the one used for investigating the *Otx* expression during *Lineus* development, we examined the expression pattern of the three *Lineus Six* genes in developing *L.viridis* larvae. We have cloned part of the homeobox of the *Lv-Six1/2* and *Lv-Six3/6* genes. Sequence comparisons with the corresponding *Ls-Six1/2* and *Ls-Six3/6* genes revealed an identity of 100% at the amino acid level, and an identity from 92% to 95% at the nucleotidic level (see appendix 6.). This suggests that these two homeobox sequences are sufficiently conserved to allow cross-hybridization of a *L. sanguineus* probe with the *L.viridis* transcript, under highly stringent conditions in *in situ* hybridization experiments. Therefore, we used RNA anti-sense probes comprising part of the sixbox and homeobox of the corresponding *L.sanguineus* gene sequences.

In one-week old larvae, *Lv-Six1/2* expression is restricted to the developing head part (Fig. 3.23). We can observe two symmetric patches of cells, which are expressing *Lv-Six1/2*. Judged from their location, these cells probably correspond to the two first eyes of the developing *Lineus* worm. In addition, *Lv-Six1/2* seems to be expressed at the level of the lateral nerve cords, which are emerging from the brain. The most anterior *Lv-Six1/2* expression observed in Fig. 3.23 corresponds to expression in anterior sensory organs, such as the frontal organ.

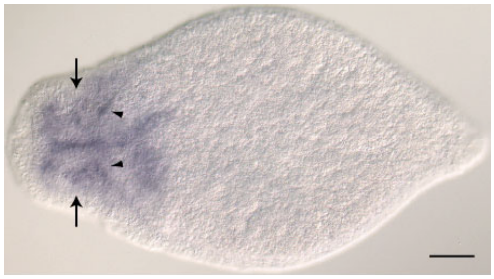


Fig. 3.23 Expression analysis of *Lv-Six1/2* in a developing *L.viridis* larva
Whole-mount ISH was done using a *Ls-Six1/2* anti-sense RNA probe. Anterior is oriented to the left, dorsal view. The larva is one-week old in accordance with the number of days post cocoon collection from our *L.viridis* laboratory stock. Arrows indicate the localization of the two first developing eyes. Arrowheads point at a possible localization of *Lv-Six1/2* expression with the emergence of the developing lateral nerve cords from the developing brain. Scale bar: 250 μm.

The expression pattern of *Ls-Six1/2* in adult *L.sanguineus* was difficult to determine, as we didn't get reproducible expression patterns. We sometimes observed staining that argues for an expression of *Ls-Six1/2* in the eye. Despite the fact that the worm, used for the whole mount ISH, harbored several eyes on each side of its head, expression of *Ls-Six1/2* is observed only at the presumptive level of a single eye (Fig. 3.24). The number of eyes is variable in adult *L.sanguineus*. In addition, the number of eyes from one side of the head, compared to the other, can be different. It is known that, during the adult life of *L.sanguineus*, new eyes can appear on the head, not necessarily in a symmetric pattern. Thus, the single patch of cells expressing *Ls-Six1/2* in the adult of Fig. 3.24 could actually correlate with the presence of a newly differentiated eye.

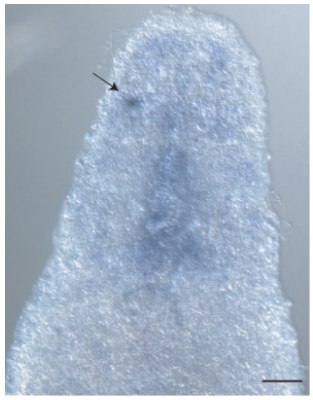


Fig. 3.24 Expression analysis of *Ls-Six1/2* in an adult *L.sanguineus* worm
Whole-mount ISH was done using a *Ls-Six1/2* anti-sense RNA probe. Only the head part has been photographed. Anterior is oriented to the top, dorsal view. The arrowheads point at a co-localization of *Ls-Six1/2* expression with one eye. Scale bar: 400 μm.

As *Ls-Six1/2* is expressed during eye development (Fig. 3.23), we took advantage of the regeneration capacities of *L.sanguineus* to test whether *Ls-Six1/2* is also expressed during regeneration of the eyes. Hence, we amputated the head of some adult worms and let the head region regenerate for various time intervals, before subjecting them to whole-mount *in situ* hybridization. *Ls-Six1/2* is strongly expressed at the level of the blastema, 10 days after the amputation (Fig. 3.25, A). At this stage of regeneration, we don't detect the presence of eyes, which are determined based on pigmentation. Hence, the two symmetric patches of cells, which are observed at the level of the blastema, on the left and on the right side, rather correspond to the development of other sense organs than to visual structures (Fig. 3.25).

No expression of *Ls-Six1/2* was detected in the regenerating brain (compared with the *Ls-Otx* expression pattern in 12 days regenerating worms, Fig. 2.12, B2). However the expression of *Ls-Six1/2* in a median stripe (Fig. 3.25, A and B) suggests that *Ls-Six1/2* is expressed at the level of the commisure, which is joining together the two cerebral ganglions.

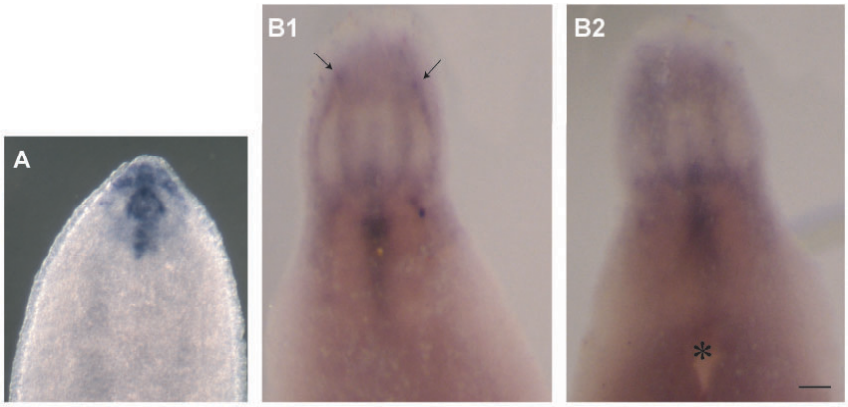


Fig. 3.25 Expression analysis of *Ls-Six1/2* during anterior head regeneration of *L.sanguineus*
Whole-mount ISH was done using a *Ls-Six1/2* anti-sense RNA probe on head amputated worms, which were undergoing regeneration. Samples were fixed 10 days (A) and one month (B1, B2) after the amputation. Anterior is oriented to the top, ventral views for A and B2, dorsal view for B. Note, in B1, the presence of two patches of stained cell, which are visible only from the dorsal side of the worm (note their absence in B2). Asterisk indicates the position of the ventral mouth in B2.
Scale bar: 1 mm.

One month after the amputation, we found *Ls-Six1/2* expressed at the level of the two newly differentiated eyes (Fig. 3.25, B1). The two spots of *Ls-Six1/2* expression are not observed from the ventral view of the animal (Fig. 3.25, B2). This is in accordance with the fact that the *Lineus* eyes are dorsally localized. In addition, the nerve fibers, which are running from the eyes to the brain, are strongly expressing *Ls-Six1/2* (Fig. 3.25, B1 and B2). A weak and diffuse staining at the periphery of the anterior region of the cerebral ganglion is also observed. This may correspond to the site where the nerve fibers emerge from the brain. Interestingly, *Ls-Otx* was expressed at a similar location, but stronger, in adult *L.sanguineus* (See Fig. 2.11, B).

3.3.3 *Ls-Six3/6* expression patterns

During development, *Lv-Six3/6* expression is restricted to the head part of the larvae. In one-week old larvae, *Lv-Six3/6* is expressed in two masses of cells, which correspond to the pair of developing cerebral ganglions (Fig. 3.26, A). In three week-old larvae, *Lv-Six3/6* is expressed in the brain, with a stronger expression in the anterior part of it (Fig. 3.26, B). It is also expressed in other body regions, anterior to the brain. These regions correspond to the cephalic gland and frontal gland, situated in the median part of the antecerebral body region. *Lv-Six3/6* is also expressed in the frontal organ, which is a sense organ, situated at the anterior tip of the worm. In addition, *Lv-Six3/6* shows a specific expression in some cells on both sides of the mouth of the larva. These clusters of cells may correspond to some glandular structures (Fig. 3.26, B).

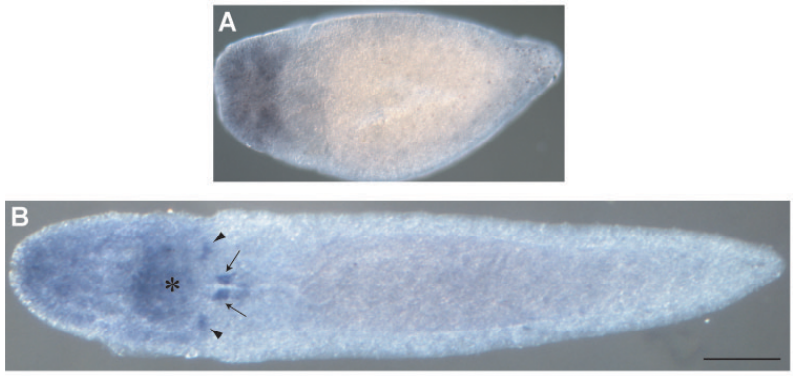


Fig. 3.26 Expression analysis of *Lv-Six3/6* in developing *L.viridis* larvae
Whole-mount ISH was done using a *Ls-Six3/6* anti-sense RNA probe. Anterior is oriented to the left, ventral view. A: one week old; B: three weeks old. Arrows point at two symmetric patches of stained cells around the mouth, while arrowheads point at the *Lv-Six3/6* expression at the level of the presumptive cerebral organ canals. The asterisk indicates strong staining in the brain. Scale bar: 450 μm.

In adult *L.sanguineus* worms, *Ls-Six3/6* does not seem to be expressed in the anterior sense organs, such as the cephalic gland, frontal gland and cephalic organ (Fig. 3.27). Some diffuse staining is observed at the level of the cerebral ganglion. This indicates that the adult brain is probably expressing *Ls-Six3/6*. However, the expression pattern observed from our whole mount ISH samples is not precise enough to determine clearly in which part of the brain it is expressed.

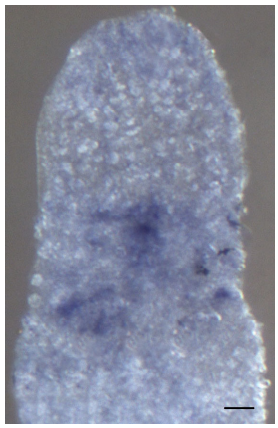


Fig. 3.27 Expression analysis of *Ls-Six3/6* in adult *L.sanguineus* worm
Whole-mount ISH was done using a *Ls-Six3/6* anti-sense RNA probe. Only the head part has been photographed. Anterior is oriented to the top, dorsal view. Scale bar: 200 μm.

3.3.4 *Ls-Six4/5* expression patterns

In 15 days-old larvae, *Lv-Six4/5* expression is restricted to the head region (Fig. 3.28, A). It is expressed in the posterior region of the developing cerebral ganglia and also in the developing sense organs that are situated just posterior to the cerebral ganglia, the cerebral organs. In addition, it is strongly expressed as two lateral stripes, running from the cerebral ganglia (=brain) to a more anterior region. When comparing the locations of the eyes in a *L.viridis* larva with the localization of the staining in the same larvae, it appears that the expression of *Lv-Six4/5*, in two lateral stripes, may correspond to expression in the nerve fibers that innervate the eyes.

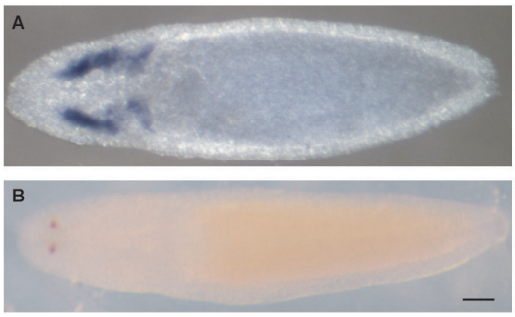


Fig. 3.28 Expression analysis of *Lv-Six4/5* in developing *L.viridis* larvae
Whole-mount ISH was done using a *Ls-Six4/5* anti-sense RNA probe. Anterior is oriented to the left, ventral view. Larvae are 15 days old. A: sample subjected to whole-mount ISH; B: same animal photographed before paraformaldehyde fixation, notice the presence of two eyes, as two pigmented spots. Scale bar: 300 μm.

Similarly to its expression during development, *Ls-Six4/5* is expressed in a posterior region of the brain and in the internal side of the cerebral organs of adult *L.sanguineus* (Fig. 3.29). A strong and precise staining is also observed at the level of structures that we previously identified as nerve fibers.

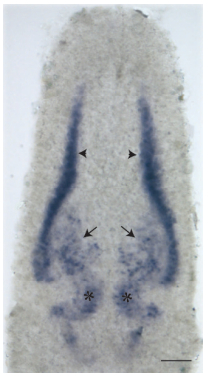


Fig. 3.29 Expression analysis of *Ls-Six4/5* in adult *L.sanguineus* worm
Whole-mount ISH was done using a *Ls-Six4/5* anti-sense RNA probe. Only the head part has been photographed. Anterior is oriented to the top, dorsal view. Arrows point at staining observed at the level of the posterior and rather external part of the dorsal cerebral ganglion, while arrowheads point at staining observed at some structures likely to be the nerve fibers that are innervating the eye structures. Asterisks indicate staining at the level of the cerebral organs. Scale bar: 450 μm.

3.4 Cross-hybridization of an antibody anti-Dachshund with *Lineus* eyes

The hunt for a *dachshund* homologue in the genome of *L.sanguineus* has not been successful, so far. Nevertheless, we can observe an antibody cross-reaction with a monoclonal *Drosophila* antibody anti-Dac: We have found a cross-reaction of the antibody with the eye structures of *L.viridis* larvae (Fig. 3.30). This is in good correlation with the expected expression pattern of a presumptive *Lineus Dachshund* homologue.

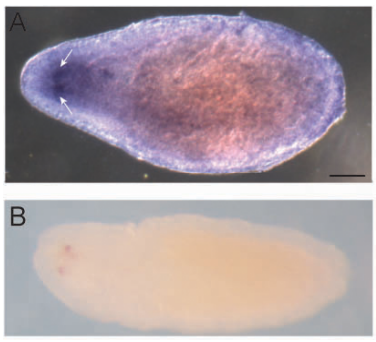


Fig. 3.30 Cross-reaction of a monoclonal antibody developed against *Drosophila* Dac with *L.viridis* larvae
A: sample that has been subjected to an immunohistology protocol; B: anesthetized larva (B), notice the presence of two eyes, as two pigmented spots. Larvae are one week old. The staining obtained by the detection of the anti-dac antibody seems to be restricted to the eyes, as indicated by the arrows. Scale bar: 200 μm.

4. Discussion

4.1 Conservation of a PST domain at the C-terminus of LsPax-6

Pax-6 transcription factors contain two highly conserved DNA-binding domains, a PD and a paired-type HD. In addition to these two DNA-binding domains, they also contain another conserved domain at their C-terminus. Indeed, the C-terminal region of Pax6 has been shown to be evolutionary highly conserved from the mollusks (Tomarev *et al.*, 1997), over the flies (Czerny *et al.*, 1999) and ascidians (Glardon *et al.*, 1997) to humans (Ton *et al.*, 1991). This domain is rich in proline, serine and threonine domain and is referred to the PST domain. It has been reported to function as a transactivation domain (Glaser *et al.*, 1994; Carriere *et al.*, 1995). The transactivation function is distributed throughout the PST domain (Tang *et al.*, 1998). The extreme sequence conservation of the PST domain is remarkable, as transactivation domains, in contrast to DNA-binding domains, are rarely conserved among transcription factors from evolutionary distant species.

Unexpectedly, we found that the published LsPax-6 sequence was not exhibiting any sequence conservation at its C-terminus (Loosli *et al.*, 1996). As this was surprising, we performed RACE PCR on *L.sanguineus* cDNA to verify whether the available sequence was correct. It appeared that the *LsPax-6* sequence was wrong and incomplete at its 3'end, due to a misassignment of an exon/ intron boundary during the analysis of the sequencing data. The correct *LsPax-6* sequence encodes a protein of 489 aa. When aligning the deduced amino acid sequence of LsPax-6 with homologues from various species, a PST domain appeared to be present in its newly identified C-terminal part. Recent secondary structure analysis has predicted the presence of two beta sheets, one in the "GLISP" motif and one in the "SVPVQ" motif, within the PST domain of the majority of Pax-6 homologues (Cooper and Hanson, 2005). Notably, the exact consensus sequence of these two motifs is found in the new LsPax-6 PST domain.

4.2 A *LsPax-6* mRNA variant encodes a truncated LsPax-6 form

In *Drosophila*, the C-terminal region of Ey and Toy differ considerably. Unlike Toy, there is less homology between the C-terminal region of vertebrate Pax-6 and Ey. Actually, the PST domain is conserved only in Toy, which seems to be more related to other Pax-6 homologues than Ey. It has been proposed that the functional differences, observed between Ey and Toy, rely, in part, on the differences in their C-termini (Punzo *et al.*, 2004). The vast majority of missense or nonsense mutations that have been associated with human diseases, such as aniridia, familial foveal dysplasia, keratitis and Peter's anomaly, are found in either the PD or the HD of the human Pax-6 (Azuma *et al.*, 1996; Hanson *et al.*, 1999; Chao *et al.*, 2000). Nevertheless, mutations, responsible for severe eye phenotypes, have also been identified in the PST domain of the human Pax-6 (Azuma and Yamada, 1998; Baum *et al.*, 1999). Hence, the proper function of Pax-6 relies not only on its DNA binding activity but also on its transactivation function. It is generally accepted that transactivation domains enhance the assembly of the transcriptional pre-initiation complex via protein-protein interactions, either directly, with components of the basal transcription apparatus, or indirectly, via additional transcription co-activators (Tjian and Maniatis, 1994).

Interestingly, we have found a variant *LsPax-6* mRNA sequence. It is generated by an alternative splicing mechanism and possesses an insertion of 14 nucleotides, which leads to a shift in the ORF, responsible for the introduction of a premature termination codon. Hence, this variant *LsPax-6* mRNA encodes a putative truncated protein that lacks the most conserved region of the PST domain: the highly evolutionary conserved motifs "GLISP" and "SVPVQ" are missing. A functional significance of such truncated LsPax-6 protein is far from being clear as we don't know whether the variant *LsPax-6* mRNA is really translated *in vivo*. Indeed, a mRNA quality-control pathway, referred as "nonsense-mediated mRNA decay", is conserved among eukaryotes. These mRNA degradation mechanisms specifically eliminate transcripts that carry premature termination codon and consequently, prevent truncated proteins from being expressed *in vivo* (Conti and Izaurralde, 2005). However, it is noteworthy that the premature stop codon, resulting from the alternative splicing of *LsPax-6*, specifically leads to the truncation of the majority of the PST domain, which has important transactivation function. Hence, such a

truncated form might not result from splicing errors but might rather have functional significance. As the transactivation potential is not localized but rather distributed throughout all the PST domain (Tang *et al.*, 1998), the hypothetical truncated LsPax-6 protein might still retain some transactivation potential. It is known that the presence of the PST domain modulates the DNA binding capacity of Pax-6 *in vitro*: substitutions of the most C-terminal amino acids have been reported to prevent the HD from binding to the DNA, whereas mutations that truncate the PST domain result in a stronger DNA-binding capacity of the truncated Pax-6 (Carriere *et al.*, 1995; Singh *et al.*, 1998; Singh *et al.*, 2001). From in *vitro* data, it appears that such truncated Pax-6 proteins compete with the full-length Pax-6 protein for the recognition of target DNA-binding sites. Indeed, these truncated proteins function as dominant-negative repressors, as they bind DNA with higher affinity but present different transactivation potential and consequently do not lead to the canonical Pax-6 activity (Singh *et al.*, 1998). An *in vivo* confirmation of the expected dominant-negative function of PST truncated form of Pax-6 proteins has been obtained from *Xenopus* studies: the injection of a mRNA encoding a PST-truncated form of *Xenopus* Pax-6 into *Xenopus* embryos has revealed that such a *Xenopus* Pax-6 isoform inhibits the formation of both, endogenous and ectopic eyes, in a dose-dependent manner (Chow *et al.*, 1999). Hence, the variant *LsPax-6* mRNA could encode either a dominant-negative repressor or an isoform with different transactivation potential. In both cases, the truncated LsPax-6 protein would probably display distinct functions than the canonical one. Alternative splicing, responsible for the generation of transcription factors with distinct or even opposite activities, switching for example from an activator to a repressor, is a well-known mechanism (Foulkes and Sassone-Corsi, 1992). Alternative splicing, which leads to isoforms with different transactivation activity, is a commonly used strategy for generating diversity among the Pax superfamily: alternative splicing of *Pax-8* transcripts have been reported to generate different Pax-8 variants, with distinct transactivation potentials, which are temporally and spatially regulated during the early development of mouse (Kozmik *et al.*, 1993). In addition, several tissue specific vertebrate *Pax7* transcripts, which encode isoforms with distinct transactivation specificities, have been identified (Lamey *et al.*, 2004). Hence, a similar mechanism might exist to generate a greater diversity of *LsPax-6* function in *L.sanguineus*. In contrast to other Pax proteins, no endogenous functional PST truncated Pax-6 proteins have been identified so far. In

fact, only one endogenous *Pax-6* splice isoform that is predicted to encode a protein lacking the PST domain has been identified in a wild-type situation. It has been isolated from a bovine eye tissue (Jaworski *et al.*, 1997). However, it is not known if this alternative splice mRNA is translated and if so, it would encode a rather unusual Pax-6 protein, as it would resemble an isolated PAI subdomain of the PD, lacking all the sequence downstream of it. Until now, no additional data have been reported from this variant bovine *Pax-6* mRNA. Our hypothetical LsPax-6 isoform, if expressed and functional *in vivo*, would be the first report of a Pax-6 protein harboring a complete PD and HD, but specifically lacking the PST domain. Based on *in situ* data, *LsPax-6* has been proposed to be involved in eye development, maintenance and regeneration, as well as in CNS development and regeneration. One can hypothesize that a temporally and spatially regulated truncated LsPax-6 form with distinct transactivation properties could serve to differentially regulate the putative numerous LsPax-6 targets during development, maintenance and regeneration of different organs. Unfortunately, as the variant *LsPax-6* mRNA sequence and the wild-type one diverged only from 14 bp, we couldn't test, by traditional ISH approach, whether there is a difference in the expression of both types of mRNA, which could have argued for distinct roles of the two proteins. One possibility would be to used, for ISH experiments, a locked nucleic acid (LNA), corresponding exactly to the anti-sense sequence of the 14 bp insert that is found in the variant *LsPax-6* mRNA sequence. This will allow us to detect specifically the expression of the variant *LsPax-6* and, then, compare it with the one known from the wild-type *LsPax-6* mRNA. LNA are commercially available oligonucleotides, which contain modified RNA nucleotides that significantly increase the thermal stability of the oligonucleotides. Due to their high melting temperatures, small LNA oligonucleotides can be used for ISH experiments and also improve standard ISH signals (Thomsen *et al.*, 2005). Indeed, LNA use is currently the only efficient method for the *in situ* detection of miRNA (Kloosterman *et al.*, 2006).

The different functional Pax-6 isoforms described so far result from alternative splicing, which occurs in the sequences coding for the DNA-binding domains. Hence, these different Pax-6 isoforms possess distinct DNA-binding properties but relatively similar transactivation potential. The situation is more complex in *Drosophila*, where the two Pax-6 homologues, Toy and Ey, possess

different transactivation properties due the highly divergent sequence of the Ey C-terminal part (Punzo *et al.*, 2004). However, these different transactivation properties of Pax-6 homologues are the results of gene duplication and subsequent divergence during evolution, and not of alternative splicing. In vertebrates, the Pax-6(5a) isoform, which results from alternative splicing, has an extra exon of 14 aa that disrupts the PAI subdomain of the PD. Thus, the use of the RED subdomain, rather than the PAI one, is responsible for the DNA-binding of the PD (Epstein *et al.*, 1994; Kozmik *et al.*, 1997). Hence, the Pax-6(5a) isoform presents a different DNA-binding capacity than the canonical Pax-6 and thus might have unique role. The ratio of these two Pax-6 forms has been shown to be critical for the normal development and function of the vertebrate eyes (Singh *et al.*, 2002). Even though a similar alternative splicing mechanism is not conserved in *Drosophila*, interestingly, Eygone (Eyg) a protein very similar to the vertebrate Pax-6(5a) has been found. Compared to Ey and Toy, Eyg is truncated and lacks part of the PAI subdomain of its PD. Like the vertebrate Pax-6(5a), Eyg binds to the "5aCON-like" sites (Jun *et al.*, 1998). The vertebrate canonical Pax-6 and the Pax-6(5a) have been proposed to be represented by functional homologues in *Drosophila*: Toy/ Ey and Eyg, respectively. A recently proposed model suggests that the *Drosophila* Eyg and the vertebrate Pax-6(5a) control tissue growth, whereas the *Drosophila* Ey and the vertebrate Pax-6 are responsible for eye tissue specification (Dominguez *et al.*, 2004). Other types of functional Pax-6 isoforms, which result from alternative splicing within the paired box, have been described (Jaworski *et al.*, 1997). For example, it is known that the *C.elegans* alternative Pax-6 isoform, mab18, which lacks the PD but retains the HD, is required for the development of the peripheral nervous system (Epstein *et al.*, 1994). Similar PD-less isoforms have been identified in other species, like in the quail and in the mouse (Carriere *et al.*, 1993; Mishra *et al.*, 2002). As they have been found in Deuterostomia and Ecdysozoa, Pax-6 isoforms that harbor a truncated or missing PD seem to be evolutionary conserved. Hence, we have searched for the presence of other variant *LsPax-6* mRNAs by RT-PCR experiments, using specific primers flanking the paired box, on mRNA extracted from various adult and regenerating tissues. However, we didn't identify any other *LsPax-6* alternative splice forms. However, it is still conceivable that other *LsPax-6* alternative splice forms are generated to achieve stage-specific function during some developmental or regenerating stages that we didn't investigate. In addition, as we used primers

flanking the paired box, it is also possible that we missed other *LsPax-6* isoforms produce from alternative splicing that occurs outside of the paired box.

4.3 Commonly accepted properties of the *Pax-2/5/8*-type HD are challenged by lophotrochozoan data

We have isolated, from *L.sanguineus* genome, a new *Pax* gene, which encodes a protein that possesses characteristic Pax-2/5/8-type amino acids in its PD. However, the assignment of this newly identified *LsPax* gene as a *LsPax-2/5/8* homologue was not obvious because it encodes a protein with a rather unusual partial HD. Based on deuterostome and ecdysozoan sequence comparisons, common Pax-2/5/8 proteins are thought to contain, in addition to a PD and an octapeptide, a partial HD, which is composed of the helices I and II, only. The structure of the deduced protein encoded by the newly identified *LsPax* gene is relatively different from the structure of common Pax-2/5/8 proteins: although it contains both Pax2/5/8-type PD and octapeptide, it seems, at a first sight, to lack a HD. More investigations have revealed that the deduced LsPax-2/5/8 protein contain some conserved residues of an HD, but surprisingly, they are from the third helix of an HD and not from the two first helices.

Some time after the isolation of the *LsPax-2/5/8* gene, another new *Pax* gene has been identified from the genome of another lophotrochozoan member, the mollusk *Arca noae*, in our laboratory (L.Keller, unpublished). Surprisingly, this *Arca Pax* gene encodes a putative protein that appears structurally similar to the one encoded by the *LsPax* -2/5/8 gene. By aligning the amino sequences of these two proteins, we found that they share an interesting common feature: both sequences do not contain residues from the two first helices of an HD, but they present the conservation of several crucial amino acids of the helix III of a HD. Furthermore, the addition of the *Platynereis* Pax-2/5/8 sequence, the only lophotrochozoan Pax-2/5/8 sequence available from databases, to our alignment revealed the conservation of the same residues of the third helix of a HD in the *Platynereis* sequence. This suggests that the conservation of these residues from the third helix of a HD might be a common feature of lophotrochozoan Pax-2/5/8 homologues. Hence, it seems that, while the partial HD, found in deuterostome and ecdysozoan Pax-2/5/8 proteins, would have retained only the two first helices of a canonical HD, the partial HD,

found in lophotrochozoan Pax-2/5/8 proteins would have retained only some of the crucial amino acids of the third helix. Until now, only three identifications of full-length *Pax-2/5/8* homologues have been reported from lophotrochozoan genomes. In fact, according to databases, some other lophotrochozoan *Pax-2/5/8* genes have been cloned, but the sequences are incomplete. Probably because their cloning was performed using degenerated primers designed against conserved sequences within the paired box, only the paired box sequences of these other *Pax-2/5/8* homologues are available. RACE PCR extension of these fragments is required in order to test whether the protein that they encode also retain residues from the third helix of a HD in their sequences. From our Pax-2/5/8 protein sequence alignment (see Fig. 3.15), it appears that the sponge proteins, which contain a Pax-2/5/8-type PD, also retain the same residues from the third helix of a HD like lophotrochozoan Pax-2/5/8 proteins. But the situation is more complicated as they also retain some residues of the two first helices of a HD (Hoshiyama *et al.*, 1998). Furthermore, the PaxB protein of *Trichoplax,* members of the putative diploblast phylum, the placozoa, harbors structural features of both Pax-2/5/8 and Pax-6 proteins. It contains a Pax-2/5/8-type PD, an octapeptide and a complete HD, close to the Pax-6-type HD (Hadrys *et al.*, 2005). Hence, the sponge Pax-2/5/8 proteins, and even more, the placozoan PaxB could represent ancestral forms from which the deuterostome/ecdysozoan-type Pax-2/5/8 proteins and the lophotrochozoan-type ones have been generated. Thus, although further investigations are required, the sequence data obtained from Lophotrochozoa might bring more light on the history of the Pax family members.

It is noteworthy that these conserved residues, which are found in all the three lophotrochozoan Pax-2/5/8-type partial HDs, are from the third helix of a HD. The third helix of the HD is indeed known to be the recognition helix crucial for specific DNA binding. The paired-class HDs are known to cooperatively bind to palindromic DNA sequences of the TAAT$(N)_{2-3}$ATTA-type, called P2 or P3 depending of the number of nucleotides that separate the two TAAT/ATTA palindromic core sequences (Wilson *et al.*, 1993). It also been shown that the third helix of paired class HDs mediates the binding to other HDs and PDs (Bruun *et al.*, 2005). The conserved residues of the third helix of a HD, which are found in lophotrochozoan-type Pax-2/5/8 partial HDs, are amino acids well known from crystallography studies to be crucial for binding. This includes a tryptophane (W) and a tyrosine (Y), which are

important aromatic amino acids that form the hydrophobic core. Hence, due to their unusual partial HD, the lophotrochozoan-type Pax-2/5/8 proteins may have very different DNA binding and protein-protein interaction properties than the deuterostome/ecdysozoan-type ones. We will investigate this possibility in the near future by firstly testing whether our LsPax-2/5/8 protein is able to bind to HD target sites, in gel-shift assays.

4.4 Identification of an unusual *Ls-Opsin* specifically expressed during eye development

We have cloned an *opsin* gene from *L.sanguineus.* The deduced Ls-Opsin amino acid sequence presents characteristic signatures, such as the K50, which allow us to clearly assign it to the opsin subfamily of GPCRs. However, the Ls-Opsin sequence does not cluster with the canonical rhabdomeric-type opsin, which is commonly found in Protostomia. As nemerteans contain rhabdomeric photoreceptors in their eyes (Vernet, 1970), the Ls-Opsin probably corresponds to a new, previously unidentified, invertebrate type of opsin. We had been looking for a *Ls-opsin* to be used as an "eye molecular-marker", since we expected it to be expressed specifically in the *Lineus* photoreceptor cells. But, the situation appeared to be more complicated: even though the cloned *ospin* is nicely expressed at the level of the visual structures of *L.viridis* larvae, no expression in adult or regenerating *L.sanguineus* eyes has been observed. Therefore, we believed that at least another type of *opsin* must be present in *L.sanguineus* genome.

It is known that only a limited range of wavelengths can be detected by each photoreceptor: this reflects its "spectral phenotype", which is determined by the type of opsin the photoreceptor contains. However, the "spectral phenotype" of one photoreceptor should not be assumed to remain unchanged throughout the animal life. In fact, several situations, in which different *opsins* are expressed during development and adult life, are known: for example, the single *opsin*-type that is expressed in the visual structures of the winter flounder is known to be replaced by three different *opsin*-types in the photoreceptors of the adults, after metamorphosis (Evans *et al.*, 1993). Similarly, when young salmon fish, living in surface waters, where ultraviolet light is abundant, move to deeper waters, where blue-green light prevails, they remodel their color vision by a switch between different opsins. Indeed, cone

photoreceptors of young salmon express only UV-*opsin*, whereas cone photoreceptors of salmon that have moved to deep waters express only blue-*opsin* (Cheng and Novales Flamarique, 2004). Furthermore, not only temporal changes, but also spatial changes in the expression of different *opsin*-types have been reported during development of zebrafish (Takechi and Kawamura, 2005). From all these findings, it appears that both, spatial and temporal expressions of different *opsin*-subtypes are more plastic than one would have thought. Thus, the isolated *Ls-opsin* is likely to be specific for some developmental stages. As we didn't find its expression during regeneration of the head regions, where new eyes are forming, the isolated *Ls-opsin* appears not to be involved in eye-regeneration. However, one could still argue for an expression in very specific stages of eye-regeneration that we didn't investigate.

We are currently still searching for other *opsin* genes in the genome of *L.sanguineus*, as it is unlikely to possess only one. As the newly identified *opsin* is not expressed in the adult eye photoreceptors, we are convinced that the *L.sanguineus* genome contains at least one other *ospin* gene, which would be expressed in the adult eye photoreceptors. This yet-unidentified *Ls-opsin* could be more similar to the canonical rhabdomeric-type *opsin*. However, as we cannot exclude the possibility that the canonical rhabdomeric-type *opsin* has been lost during evolution by the *L.sanguineus* genome, the other unidentified *Ls-opsin* could also be from an unusual invertebrate-type. In addition to the rhabdomeric photoreceptors found in inverse adult nemertean eyecups (Vernet, 1970), some ciliary photoreceptors have been described in nemertean brains (Vernet, 1974). Hence, it is also conceivable that a ciliary-type *opsin* could be isolated from *L.sanguineus*, as it was recently achieved for the marine rag-worm *Platynereis*, which also have some ciliary photoreceptors located in their brain (Arendt *et al.*, 2004).

4.5 First identification of homologues from the three *Six* gene subfamilies in a lophotrochozoan genome

From their basal placement in phylogenetic trees (Agosti *et al.*, 1996), a greater antiquity of the *Six* gene family, compared to other homeobox-containing gene families, has been proposed. *Six* homologues have already been isolated throughout basal Metazoa, including several sponges, cnidaria and a ctenophore (Bebenek *et al.*,

2004), making the presence of *Six* homologues in the genome of *L.sanguineus* likely. Indeed, we have identified three different *Ls-Six* genes, one of each *Six* gene subfamily: *Ls-Six1/2, Ls-Six3/6* and *Ls-Six4/5*.

Chromosomal linkages of different *Six* genes have been reported for several bilaterian species. From ecdysozoan genome studies, we know that *so* and *optix* map very closely on chromosome 2R in *Drosophila* (Toy *et al.*, 1998). Also the four *C.elegans Six* genes are located on a same chromosome (Ruvkun and Hobert, 1998). The situation is similar in deuterostomian genomes: the *Six2* and *Six3* genes map on the same mouse chromosome (Oliver *et al.*, 1995a; Oliver *et al.*, 1995b) and the *Six1* and *Six4* and *Six6* genes, which are members of the three different *Six* subfamilies, are closely linked on the same human chromosome (Gallardo *et al.*, 1999). Based on these chromosomal linkage findings and on the fact that ecdysozoan and deuterostome genomes retain at least one gene from each of the three *Six* subfamilies, it has been proposed that the genome of the Urbilateria possessed only three *Six* genes, which were arranged in a single cluster (Gallardo *et al.*, 1999). This cluster supposedly resulted from two successive tandem duplications of a single gene precursor. As we isolated the first lophotrochozoan *Ls-Six4/5*, our work on *Ls-Six* homologues corresponds to the first report of the presence of one gene from each of the three *Six* subfamilies in a lophotrochozoan genome. Despite our extensive PCR survey, we could isolate only one gene from each *Six* subfamily. This fact suggests that *Ls-Six* genes could have retained some features of the *Six* genes that were present in the urbilaterian genome. To corroborate this idea, we have thought to look in a near future for their genomic organization to check whether the *Ls-Six* genes are arranged into a cluster, similarly to the hypothetical archetypal one from the Urbilateria.

4.6 General involvement of the three *Ls-Six* genes in CNS and anterior sensory organ development

According to its expression pattern, *Ls-Six1/2* is likely to be involved in both, development and regeneration of anterior *Lineus* sense organs, including the eyes. It seems involved, as well, in the development and regeneration of some part of the brain. It is known that of *Six1/2* homologues display conserved functions in invertebrate eyes. Indeed, in collaboration with *ey*, *eya* and *dac*, *so* is required for the

Drosophila compound eye development (Bonini *et al.*, 1997; Shen and Mardon, 1997; Halder *et al.*, 1998). In fact, *so* plays a crucial role for the development of the entire *Drosophila* visual system, including not only the adult compound eyes but also, the embryonic optic lobe, the Bolwig's organ, which is the larval photoreceptor organ, and the adult ocelli (Halder *et al.*, 1998; Niimi *et al.*, 1999; Punzo *et al.*, 2002). Similarly, *Six1/2* homologues are also expressed in the visual system of Lophotrochozoa, such as polychaete and planarian (Pineda *et al.*, 2000; Arendt *et al.*, 2002). Furthermore, the inactivation of the planarian *G.tigrina Six1/2* homologue, *Gtso*, by RNAi blocks eye regeneration completely (Pineda *et al.*, 2000). Likewise, we have also argued for a role of *Ls-Six1/2* during regeneration of *Lineus* eyes. However, in contrast to *Ls-Six1/2*, the planarian *Gtso* remains expressed in fully regenerated eyes (Pineda *et al.*, 2000). Hence, *Ls-Six1/2* does not seem important for the maintenance of the *Lineus* adult eyes. As the *Six1/2* homologues present conserved expression patterns in both, ecdysozoan and lophotrochozoan species, their roles in early visual system specification have been proposed to be evolutionarily ancient. This idea has been reinforced by the expression of a *Six1/2* homologue in the cnidarian eye cup (Stierwald *et al.*, 2004), which suggests a conservation of this role outside the Bilateria. However, this view is challenged by vertebrate expression data. The vertebrate *Six1* and *Six2* do not participate to the early eye development. They are only involved in the late differentiation of the vertebrate retina (Kawakami *et al.*, 1996; Ghanbari *et al.*, 2001). Thus, the early development of visual structures differs among bilaterian animals, as the involvement of *Six1/2* gene in eye development is conserved throughout Protostomia but not in Deuterostomia. Other roles, such as involvement in myogenesis and in early development of numerous organs like the kidney and the thymus, have also been reported for the vertebrate *Six1* and *Six2* genes (Laclef *et al.*, 2003a; Laclef *et al.*, 2003b; Stierwald *et al.*, 2004; Grifone *et al.*, 2005). In addition, they also are important for the patterning of the head and the development of many sensory organs, such as the olfactory system and the auditory system (Ghanbari *et al.*, 2001; Laclef *et al.*, 2003; Zheng *et al.*, 2003). Interestingly, *Ls-Six1/2* is also specifically expressed at the level of several *Lineus* sense organs, not only in the eyes but also in the frontal gland and in the frontal organ, during both, their development and their regeneration. This suggests a conservation of *Six1/2* homologue function in sensory organ development. In addition, the strong and broad *Ls-Six1/2* expression during early head regeneration suggests the conservation of

broader head patterning function of *Six1/2* genes in *Lineus*. Furthermore, during development and regeneration, *Ls-Six1/2* is expressed in specific regions of the brain, namely at the level of the commisure that is joining together the two cerebral ganglions and at the periphery of the anterior region of the cerebral ganglia, which probably corresponds to the site where nerve fibers emerge from the brain to innervate the eyes. Our expression data argue for an involvement of *Ls-Six1/2* in the development and regeneration of some specific CNS regions of *Lineus*, even though, in contrast to *Six3/6* subfamily, no function of *Six1/2* subfamily has been proposed to be evolutionary conserved in CNS development.

Six3 and *Six6* genes demarcate almost the entire vertebrate forebrain region, which is composed of the telencephalon and the diencephalon. Studies from mouse, chicken, frog and fish argue for an early expression of the vertebrate *Six3* and *Six6* genes in the anterior-most part of the neural plate. Their expressions are later maintained in the CNS, in the visual system, in broad regions of the hypothalamus, in the pineal, in the ventral telencephalon and in the thalamus (Oliver *et al.*, 1995a; Bovolenta *et al.*, 1998; Loosli *et al.*, 1998; Jean *et al.*, 1999; Ghanbari *et al.*, 2001). Similarly, *Six3/6* subfamily members are also expressed in the anterior-most region of protostomian brains (Pineda and Salo, 2002a; Tessmar-Raible, 2004). Hence, *Six3/6* homologues seem to display evolutionary conserved roles in the development of the CNS. In accordance with this hypothesis, we have found *Ls-Six3/6* to be expressed in the developing *Lineus* brain, with the strongest expression at the most anterior part of the cerebral ganglia. Furthermore, it also seems to remain expressed in the adult brain. Based on studies from vertebrates, *Drosophila* and even cnidarian, an additional evolutionary conserved role of the *Six3/6* homologues has been proposed. *Six3* and *Six6* homologues are expressed in the developing eyes of the mouse embryo (Oliver *et al.*, 1995a; Toy and Sundin, 1999). Furthermore, enlargement of the optic stalk is observed when the *Six3* homologue is ectopically expressed in zebrafish (Kobayashi *et al.*, 1998). In addition, *Six3* over-expression in medaka fish leads to the formation of an ectopic retina, while over-expression of the *Six6* homologue in the *Xenopus* results in increased eye size (Loosli *et al.*, 1999; Zuber *et al.*, 1999). Similarly, *optix*, the *Drosophila Six3/6* homologue, has been reported to be able to induce ectopic eyes, independently of *ey* (Seimiya and Gehring, 2000). Surprisingly, our expression data for *Ls-Six3/6* do not suggest, as one could have expected, any involvement of this

Six3/6 homologue neither in the development nor in the maintenance of the *Lineus* eyes. Strikingly, *Gtsix-3*, the planaria *Six3/6* homologue is expressed in brain branches but is never detected in eyes, neither in differentiated adult eyes nor in regenerating ones (Pineda and Salo, 2002a). Hence, although *Six3/6* homologues are clearly involved in the development of deuterostome and ecdysozoan eyes, it does not seem to be the case for the development of the lophotrochozoan eyes, so far. Thus, further investigations of lophotrochozoan *Six3/6* homologues are required for the validation of the assumed conserved role of these genes in the development of bilaterian visual structures.

An evolutionary conserved function of *Six4/5* homologues in mesoderm patterning has been proposed. Indeed, mutations in the *C.elegans Six5* homologue, *unc-39*, lead to defects in specification and differentiation of the mesoderm, in addition to defects in neuronal migration and axon pathfindings (Yanowitz *et al.*, 2004). Likewise, *D-six4* is required for the development of most *Drosophila* cell types originating from the non-dorsal mesoderm, such as the fat body, somatic cells of the gonad and specific subset of muscles. Furthermore, the misexpression of *D-six4* and its partner, *eya*, is sufficient to impose a "non-dorsal mesodermal" fate on other mesodermal cells, highlighting the role of *D-six4* as a key mesodermal patterning mediator (Clark *et al.*, 2006). In addition to the co-expression of *Six4* and *Six5* during vertebrate myogenesis, an early involvement of *Six4* during mesoderm development has been reported in vertebrates (Grifone *et al.*, 2005). Hence, data from Deuterostomia and Ecdysozoa argue for a conserved role in mesoderm patterning and for a more broad function in cell motility and differentiation. From our expression data, *Ls-Six4/5* expression is restricted to the head region, where it shows similar expression in both, developing larvae and adult worms. In respect with its expression pattern, we suggest a role of *Ls-Six4/5* in the innervation of sense organs. It is indeed strongly expressed in nerve fibers that run from the brain to the anterior sense organs, including the eyes. In addition, it is expressed in the cerebral organs, which are also sense organs, and is expressed too in the regions of the brain that is connected to these cerebral organs. According to the databases, we have reported the isolation of the first lophotrochozoan *Six4/5* homologue. Based on its expression pattern, we have suggested an involvement in the development and the innervation of the sense organs.

To which extend this *Six4/5* presumed function is conserved will remain unclear as long as additional data from lophotrochozoan homologue are not available.

4.7 Some members of the Retinal Determination Genetic Network are co-expressed in developing, regenerating and differentiated *Lineus* eyes

The *L.sanguineus Pax-6* homologue, *LsPax-6*, has been shown to be expressed in a rather broad region of the head, including the eye region (Loosli *et al.*, 1996). In addition to its expression in the regenerating brain, *LsPax-6* is also specifically detected in newly differentiating eyes (Loosli *et al.*, 1996). Furthermore, we have previously reported that inactivation of *LsPax-6* by RNAi, mediated by repeated injections of *LsPax-6* dsRNA, results in disappearance of adult eyes and blocks the regeneration process of the eyes (Charpignon, 2002). Altogether these data strongly argue for a crucial role of *Pax-6* in both, maintenance and regeneration of the *Lineus* eyes. Our results from *LsPax-6* studies are in good agreement with the model derived from studies in vertebrates, which suggests an involvement of *Pax-6* in the maintenance and the regeneration of the retina. The vertebrate adult visual structures seem to retain the expression of *Pax-6* throughout their existence: a continuous expression of *Pax-6* has been detected in the retina throughout the lifespan of human (Stanescu *et al.*, 2005). In addition, *Pax-6* remains strongly expressed in the surface epithelia of the adult cornea and conjunctiva in chicken, mouse and monkey (Koroma *et al.*, 1997). This has led to the proposition of a direct role of vertebrate *Pax-6* in the maintenance and in the proliferation of corneal stem cells (Koroma *et al.*, 1997). Furthermore, a reduced level of *Pax-6* transcription, as a consequence of a decrease in Notch activity, has been suggested to be responsible for human retinal degeneration (Pauli, 2004). In the light of this hypothesis, it is noteworthy that the adult eyes disappear as a result of *LsPax-6* inactivation by RNAi and even more remarkable that eyes reappear once the *LsPax-6* dsRNA injections are suspended (Charpignon, 2002). Taken together, data from vertebrate *Pax-6* and *LsPax-6* studies favorably argue for an evolutionary conserved function of *Pax-6* in the maintenance of both vertebrate and invertebrate visual structures.

We wanted to further characterize the components of the *Lineus* eye specification network. Members of the Pax-6, Eya, Six and Dac families of transcription factors are known to collectively direct the formation of the invertebrate eyes through a complicated network of regulatory interactions, involving feedback regulatory loops and protein-protein interactions. Hence, we thought to search for *Lineus* homologues of these transcription factors. We have indeed isolated a *Six1/2* homologue, *Ls-Six1/2*. According to its expression pattern, *Ls-Six1/2* could be involved in the development of the eyes and also in their regeneration. Our data are in agreement with the ones obtained from the expression patterns of other lophotrochozoan *Six1/2* homologues. Indeed, *Six1/2* homologues from different species of planaria are expressed in regenerating eyes and have been shown to be critical for their regeneration (Pineda *et al.*, 2000; Mannini *et al.*, 2004). But, in contrast to *Ls-Six1/2*, the *Six1/2* homologue remains expressed in the fully developed eyes of planarians (Pineda *et al.*, 2000; Mannini *et al.*, 2004). Since no *Ls-Six1/2* expression was observed in adults, we have concluded that its expression is not required for the maintenance of the adult *Lineus* eyes. However, it is possible that *Ls-Six1/2* is expressed at a lower level in adults for the whole mount ISH detection, similar as it was found for the two planarian *Pax-6* homologues: the planarian *Pax-6A* and *Pax-6B* transcripts can be detected in both regenerating and adult eyes only by electron microscopy *in situ* technique (Callaerts *et al.*, 1999; Pineda *et al.*, 2002b). Surprisingly, twice we have observed, in adult worms, a very localized *Ls-Six1/2* expression that was likely co-localizing with an eye. As such expression pattern was detected only twice over all our ISH experiments and as visual structures are known to continuously form during the life of *Lineus* worm, we propose that this transient *Ls-Six1/2* expression could be present due to the differentiation of a new eye in an intact adult. This would be in good agreement with the idea that the expression of *Ls-Six1/2*, similarly to the one of *LsPax-6*, may be essential for the formation of an eye in *Lineus.* However, more investigations are required to validate this finding. Indeed, although a *Pax-6* involvement in planarian eye regeneration and maintenance is suggested based on the expression patterns of the planarian *Pax-6* homologues, astonishingly the inactivation by RNAi of both *Pax-6A* and *Pax-6B* seems to inhibit neither eye regeneration nor eye maintenance (Pineda *et al.*, 2002b). We have planned to test RNAi on *Ls-Six1/2* in head regenerating worms. On one hand, this will allow us to assess the *Ls-Six1/2* function during regeneration of the eyes. On the other hand,

by checking for *LsPax-6* expression in such *Ls-Six1/2* dsRNA treated worms, we will be able to investigate the presumptive epistatic relations between *Ls-Six1/2* and *LsPax-6.* Reciprocally, it is important to repeat the *LsPax-6* inactivation experiment in order to search for a hypothetical remaining *Ls-Six1/2* expression, in the absence of the *LsPax-6* one.

From the RNAi-induced *LsPax-6* phenotypes, we already know that *Ls-Six1/2* alone, assuming it is involved in this process, cannot trigger eye regeneration. Interestingly, the inactivation of planarian *Pax-6* homologues by RNAi seems to have no effect on eye maintenance or regeneration and no effect on the expression of *Gtso* (Pineda *et al.*, 2002b). In contrast, the inactivation of *Gtso* by RNAi leads to a no-eye phenotype in head regenerating worms and to a decrease of the eye size in non-regenerating worms (Pineda *et al.*, 2000; Pineda *et al.*, 2002b; Salo *et al.*, 2002). Hence, while the maintenance of *Lineus* eyes seems to be dependent on *LsPax-6* expression but not on *Ls-Six1/2* expression (Charpignon, 2002; this PhD work), the maintenance of planarian eyes seems to be dependent on *Six1/2* homologue expression but not on *Pax-6* homologue expression (Pineda *et al.*, 2000). Similarly, while the regeneration of *Lineus* eyes is *Pax-6* dependent, the regeneration of planarian eyes does not seem to rely on *Pax-6* activity (Callaerts *et al.*, 1999; Charpignon, 2002; Pineda *et al.*, 2002b). This highlights the fact that some specie specificities probably exist in the regulation and function of members of the evolutionary conserved RDGN, involved in invertebrate eye development. Interestingly, the embryonic and adult vertebrate eyes also express many of the RDGN members, suggesting a certain degree of conservation of the eye specification network between vertebrates and invertebrates (Hanson, 2001; Kumar and Moses, 2001a; Donner and Maas, 2004; Silver and Rebay, 2005). However from vertebrate gene studies, it appears that not only the regulation and function of a same RDGN member can vary between invertebrates and vertebrates, but also which kind of RDGN members is involved in eye specification vary between invertebrates and vertebrates. Indeed, while *Six1/2* and *Six3/6* homologues, respectively *so* and *optix*, are expressed in the developing *Drosophila* eye, only *Six3* and *Six6* genes, which belong to another *Six* family, are expressed in the developing vertebrate eye (Toy *et al.*, 1998; Kawakami *et al.*, 2000). In addition, through complex transcriptional regulation, which leads to the generation of several functional isoforms, the vertebrate

Pax-6 is able to achieve the numerous functions that are carried out by *ey, toy* and *eyg* in *Drosophila* (Dominguez *et al.*, 2004; Rodrigues and Moses, 2004). Furthermore, although from their expression data, a role in eye development has been suggested for the three vertebrate *Eya* homologues (Xu *et al.*, 1997), it appeared that the disruption of both mouse *Eya1* and *Eya2* do not affect the eye development (null mutation of the Eya3 has not yet been reported) (Purcell, 2002; Donner and Maas, 2004). Hence, the vertebrate *Eya* homologues do not appear to be as crucial for eye development as their invertebrate homologues.

During head regeneration, *LsPax-6* is expressed in a relatively broad region, where new eyes will regenerate, while *Ls-Six1/2* expression is specifically restricted to the location of the regenerating eyes. Similar observations are made for the expression pattern of both genes during *Lineus* eye development. Hence, we suggest the hypothesis that during development and regeneration, *LsPax-6* expression define a broad region that have the potential to develop *Lineus* visual structures and that subsequently, eyes will emerge only where *Ls-Six1/2* is expressed, within the *LsPax-6* expressing head territory. This hypothesis is rather speculative and probably represents only part of the story. Actually, we think that *Ls-Six1/2* expression is probably not sufficient to specify *Lineus* eye development and probably needs to be combined with the expression of other partners. As the *Lineus* eye tissues cross-react with an antibody against Dac, we have suggested the existence of a presumptive *Ls-Dach*, which is a good candidate for being an *Ls-Six1/2* interacting partner during the *Lineus* eye specification. In addition, we are also searching for an *L.sanguineus Eya* homologue, which could also be an interacting partner as, from planarian studies, the essential role of the conserved interaction between Eya and Six1/2 homologues has been recently extended to lophotrochozoan members (Mannini *et al.*, 2004). We want to find evidence whether the acquisition of "eye competence" in *Lineus* could resembles the model proposed for the "eye competence" acquisition by the cells from the eye imaginal disc: although *so*, *eya* and *dac* are expressed elsewhere in the fly embryo, they are all first co-expressed in the eye disc during the second instar larvae. It has been proposed that this co-expression, during second instar larva, is locking-in the eye competence in the cells that are expressing *ey* (Kumar and Moses, 2001b; Pichaud *et al.*, 2001). Moreover, although inductive and patterning events prepare the anterior neural plate for the formation of the vertebrate eye field, this is the

coordinated expression of several transcription factors, such as *Pax-6, Six3, Six6,* that specify the eye field (Zuber *et al.*, 2003). A similar model has been formulated for the development of the larval eye of another lophotrochozoan member. Indeed, in *Platynereis*, the larval eye precursors are formed at the intersection of the *Pax-6* and *Six1/2* territories (Arendt *et al.*, 2002).

5. References of chapter III

Agosti, D., Jacobs, D. and DeSalle, R. (1996). "On combining protein sequences and nucleic acid sequences in phylogenetic analysis: the homeobox protein case." *Cladistics* **12**: 65-82.

Arendt, D. and Wittbrodt, J. (2001). "Reconstructing the eyes of Urbilateria." *Philos Trans R Soc Lond B Biol Sci* **356**(1414): 1545-1563.

Arendt, D., Tessmar, K., de Campos-Baptista, M. I., Dorresteijn, A. and Wittbrodt, J. (2002). "Development of pigment-cup eyes in the polychaete *Platynereis dumerilii* and evolutionary conservation of larval eyes in Bilateria." *Development* **129**(5): 1143-1154.

Arendt, D. (2003). "Evolution of eyes and photoreceptor cell types." *Int J Dev Biol* **47**(7-8): 563-571.

Arendt, D., Tessmar-Raible, K., Snyman, H., Dorresteijn, A. W. and Wittbrodt, J. (2004). "Ciliary photoreceptors with a vertebrate-type opsin in an invertebrate brain." *Science* **306**(5697): 869-871.

Arthur, W. (2002). "The emerging conceptual framework of evolutionary developmental biology." *Nature* **415**(6873): 757-764.

Azuma, N., Nishina, S., Yanagisawa, H., Okuyama, T. and Yamada, M. (1996). "PAX6 missense mutation in isolated foveal hypoplasia." *Nat Genet* **13**(2): 141-142.

Azuma, N. and Yamada, M. (1998). "Missense mutation at the C terminus of the *PAX6* gene in ocular anterior segment anomalies." *Invest Ophthalmol Vis Sci* **39**(5): 828-830.

Baonza, A., Murawsky, C. M., Travers, A. A. and Freeman, M. (2002). "Pointed and Tramtrack69 establish an EGFR-dependent transcriptional switch to regulate mitosis." *Nat Cell Biol* **4**(12): 976-980.

Baum, L., Pang, C. P., Fan, D. S., Poon, P. M., Leung, Y. F., Chua, J. K. and Lam, D. S. (1999). "Run-on mutation and three novel nonsense mutations identified in the *PAX6* gene in patients with aniridia." *Hum Mutat* **14**(3): 272-273.

Bebenek, I. G., Gates, R. D., Morris, J., Hartenstein, V. and Jacobs, D. K. (2004). "*sine oculis* in basal Metazoa." *Dev Genes Evol* **214**(7): 342-351.

Bonini, N. M., Leiserson, W. M. and Benzer, S. (1993). "The *eyes absent* gene: genetic control of cell survival and differentiation in the developing *Drosophila* eye." *Cell* **72**(3): 379-395.

Bonini, N. M., Bui, Q. T., Gray-Board, G. L. and Warrick, J. M. (1997). "The *Drosophila eyes absent* gene directs ectopic eye formation in a pathway conserved between flies and vertebrates." *Development* **124**(23): 4819-4826.

Bopp, D., Burri, M., Baumgartner, S., Frigerio, G. and Noll, M. (1986). "Conservation of a large protein domain in the segmentation gene *paired* and in functionally related genes of *Drosophila*." *Cell* **47**(6): 1033-1040.

Bovolenta, P., Mallamaci, A., Puelles, L. and Boncinelli, E. (1998). "Expression pattern of *cSix3*, a member of the Six/sine oculis family of transcription factors." *Mech Dev* **70**(1-2): 201-203.

Bruun, J. A., Thomassen, E. I., Kristiansen, K., Tylden, G., Holm, T., Mikkola, I., Bjorkoy, G. and Johansen, T. (2005). "The third helix of the homeodomain of paired class homeodomain proteins acts as a recognition helix both for DNA and protein interactions." *Nucleic Acids Res* **33**(8): 2661-2675.

Bui, Q. T., Zimmerman, J. E., Liu, H. and Bonini, N. M. (2000). "Molecular analysis of *Drosophila* eyes absent mutants reveals features of the conserved Eya domain." *Genetics* **155**(2): 709-720.

Burri, M., Tromvoukis, Y., Bopp, D., Frigerio, G. and Noll, M. (1989). "Conservation of the paired domain in metazoans and its structure in three isolated human genes." *Embo J* **8**(4): 1183-1190.

Callaerts, P., Munoz-Marmol, A. M., Glardon, S., Castillo, E., Sun, H., Li, W. H., Gehring, W. J. and Salo, E. (1999). "Isolation and expression of a *Pax-6* gene in the regenerating and intact Planarian *Dugesia(G)tigrina*." *Proc Natl Acad Sci U S A* **96**(2): 558-563.

Carl, M., Loosli, F. and Wittbrodt, J. (2002). "*Six3* inactivation reveals its essential role for the formation and patterning of the vertebrate eye." *Development* **129**(17): 4057-4063.

Carriere, C., Plaza, S., Martin, P., Quatannens, B., Bailly, M., Stehelin, D. and Saule, S. (1993). "Characterization of quail Pax-6 (Pax-QNR) proteins expressed in the neuroretina." *Mol Cell Biol* **13**(12): 7257-7266.

Carriere, C., Plaza, S., Caboche, J., Dozier, C., Bailly, M., Martin, P. and Saule, S. (1995). "Nuclear localization signals, DNA binding, and transactivation properties of quail Pax-6 (Pax-QNR) isoforms." *Cell Growth Differ* **6**(12): 1531-1540.

C.elegans Sequencing Consortium (1998). "Genome sequence of the nematode *C.elegans*: a platform for investigating biology." *Science* **282**(5396): 2012-2018

Chao, L. Y., Huff, V., Strong, L. C. and Saunders, G. F. (2000). "Mutation in the *PAX6* gene in twenty patients with aniridia." *Hum Mutat* **15**(4): 332-339.

Charpignon, V. (2002). Approche fonctionnelle de *Pax-6* chez le nemertien *Lineus sanguineus* par la technique d'ARN interférence, Bâle, Reims. Diploma thesis.

Chen, R., Amoui, M., Zhang, Z. and Mardon, G. (1997). "Dachshund and eyes absent proteins form a complex and function synergistically to induce ectopic eye development in *Drosophila*." *Cell* **91**(7): 893-903.

Cheng, C. L. and Novales Flamarique, I. (2004). "*Opsin* expression: new mechanism for modulating colour vision." *Nature* **428**(6980): 279.

Cheyette, B. N., Green, P. J., Martin, K., Garren, H., Hartenstein, V. and Zipursky, S. L. (1994). "The *Drosophila sine oculis* locus encodes a homeodomain-containing protein required for the development of the entire visual system." *Neuron* **12**(5): 977-996.

Chi, N. and Epstein, J. A. (2002). "Getting your Pax straight: Pax proteins in development and disease." *Trends Genet* **18**(1): 41-47.

Chisholm, A. D. and Horvitz, H. R. (1995). "Patterning of the *Caenorhabditis elegans* head region by the *Pax-6* family member *vab-3*." *Nature* **377**(6544): 52-55.

Chow, R. L., Altmann, C. R., Lang, R. A. and Hemmati-Brivanlou, A. (1999). "*Pax6* induces ectopic eyes in a vertebrate." *Development* **126**(19): 4213-4222.

Clark, I. B., Boyd, J., Hamilton, G., Finnegan, D. J. and Jarman, A. P. (2006). "*D-six4* plays a key role in patterning cell identities deriving from the *Drosophil*a mesoderm." *Dev Biol* **294**(1): 220-231.

Conte, I., Morcillo, J. and Bovolenta, P. (2005). "Comparative analysis of *Six 3* and *Six 6* distribution in the developing and adult mouse brain." *Dev Dyn* **234**(3): 718-725.

Conti, E. and Izaurralde, E. (2005). "Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species." *Curr Opin Cell Biol* **17**(3): 316-325.

Cooper, S. T. and Hanson, I. M. (2005). "A screen for proteins that interact with PAX6: C-terminal mutations disrupt interaction with HOMER3, DNCL1 and TRIM11." *BMC Genet* **6**: 43.

Cvekl, A. and Piatigorsky, J. (1996). "Lens development and crystallin gene expression: many roles for Pax-6." *Bioessays* **18**(8): 621-630.

Czerny, T., Schaffner, G. and Busslinger, M. (1993). "DNA sequence recognition by Pax proteins: bipartite structure of the paired domain and its binding site." *Genes Dev* **7**(10): 2048-2061.

Czerny, T., Halder, G., Kloter, U., Souabni, A., Gehring, W. J. and Busslinger, M. (1999). "*twin of eyeless*, a second *Pax-6* gene of *Drosophila*, acts upstream of *eyeless* in the control of eye development." *Mol Cell* **3**(3): 297-307.

Dacey, D. M., Liao, H. W., Peterson, B. B., Robinson, F. R., Smith, V. C., Pokorny, J., Yau, K. W. and Gamlin, P. D. (2005). "Melanopsin-expressing ganglion cells in primate retina signal colour and irradiance and project to the LGN." *Nature* **433**(7027): 749-754.

Davis, R. J., Shen, W., Sandler, Y. I., Amoui, M., Purcell, P., Maas, R., Ou, C. N., Vogel, H., Beaudet, A. L. and Mardon, G. (2001). "Dach1 mutant mice bear no gross abnormalities in eye, limb, and brain development and exhibit postnatal lethality." *Mol Cell Biol* **21**(5): 1484-1490.

Dominguez, M., Ferres-Marco, D., Gutierrez-Avino, F. J., Speicher, S. A. and Beneyto, M. (2004). "Growth and specification of the eye are controlled independently by Eyegone and Eyeless in *Drosophila melanogaster*." *Nat Genet* **36**(1): 31-39.

Dominguez, M. and Casares, F. (2005). "Organ specification-growth control connection: new in-sights from the *Drosophila* eye-antennal disc." *Dev Dyn* **232**(3): 673-684.

Donner, A. L. and Maas, R. L. (2004). "Conservation and non-conservation of genetic pathways in eye specification." *Int J Dev Biol* **48**(8-9): 743-753.

Eberhard, D., Jimenez, G., Heavey, B. and Busslinger, M. (2000). "Transcriptional repression by Pax5 (BSAP) through interaction with corepressors of the Groucho family." *Embo J* **19**(10): 2292-2303.

Epstein, J. A., Glaser, T., Cai, J., Jepeal, L., Walton, D. S. and Maas, R. L. (1994). "Two independent and interactive DNA-binding subdomains of the Pax6 paired domain are regulated by alternative splicing." *Genes Dev* **8**(17): 2022-2034.

Evans, B. I., Harosi, F. I. and Fernald, R. D. (1993). "Photoreceptor spectral absorbance in larval and adult winter flounder (*Pseudopleuronectes americanus*)." *Vis Neurosci* **10**(6): 1065-1071.

Fernald, R. D. (1988). Aquatic adaptations in fish eyes. In: Sensory Biology of Aquatic Animals. New York: Springer Verlag.

Fernald, R. D. (2000). "Evolution of eyes." *Curr Opin Neurobiol* **10**(4): 444-450.

Foulkes, N. S. and Sassone-Corsi, P. (1992). "More is better: activators and repressors from the same gene." *Cell* **68**(3): 411-414.

Gallardo, M. E., Lopez-Rios, J., Fernaud-Espinosa, I., Granadino, B., Sanz, R., Ramos, C., Ayuso, C., Seller, M. J., Brunner, H. G., Bovolenta, P. and Rodriguez de Cordoba, S. (1999). "Genomic cloning and characterization of the human homeobox gene *SIX6* reveals a cluster of *SIX* genes in chromosome 14 and associates SIX6 hemizygosity with bilateral anophthalmia and pituitary anomalies." *Genomics* **61**(1): 82-91.

Gehring, W. J. and Ikeo, K. (1999). "Pax 6: mastering eye morphogenesis and eye evolution." *Trends Genet* **15**(9): 371-377.

Gehring, W. J. (2004). "Historical perspective on the development and evolution of eyes and photoreceptors." *Int J Dev Biol* **48**(8-9): 707-717.

Gehring, W. J. (2005). "New perspectives on eye development and the evolution of eyes and photoreceptors." *J Hered* **96**(3): 171-184.

Ghanbari, H., Seo, H. C., Fjose, A. and Brandli, A. W. (2001). "Molecular cloning and embryonic expression of *Xenopus Six* homeobox genes." *Mech Dev* **101**(1-2): 271-277.

Glardon, S., Callaerts, P., Halder, G. and Gehring, W. J. (1997). "Conservation of *Pax-6* in a lower chordate, the ascidian *Phallusia mammillata*." *Development* **124**(4): 817-825.

Glaser, T., Jepeal, L., Edwards, J. G., Young, S. R., Favor, J. and Maas, R. L. (1994). "*PAX6* gene dosage effect in a family with congenital cataracts, aniridia, anophthalmia and central nervous system defects." *Nat Genet* **7**(4): 463-471.

Greuet, C. (1965). Structure fine de l'ocelle d'*Erythropsis pavillardi* Hertwig, pteridinien Warnowiidae Lindemann., *C.R. Acad. Sci. (Paris)*.

Grifone, R., Demignon, J., Houbron, C., Souil, E., Niro, C., Seller, M. J., Hamard, G. and Maire, P. (2005). "Six1 and Six4 homeoproteins are required for *Pax3* and *Mrf* expression during myogenesis in the mouse embryo." *Development* **132**(9): 2235-2249.

Hadrys, T., DeSalle, R., Sagasser, S., Fischer, N. and Schierwater, B. (2005). "The Trichoplax *PaxB* gene: a putative Proto-*PaxA/B/C* gene predating the origin of nerve and sensory cells." *Mol Biol Evol* **22**(7): 1569-1578.

Halder, G., Callaerts, P. and Gehring, W. J. (1995). "Induction of ectopic eyes by targeted expression of the *eyeless* gene in *Drosophila*." *Science* **267**(5205): 1788-1792.

Halder, G., Callaerts, P., Flister, S., Walldorf, U., Kloter, U. and Gehring, W. J. (1998). "Eyeless initiates the expression of both *sine oculis* and *eyes absent* during *Drosophila* compound eye development." *Development* **125**(12): 2181-2191.

Hanson, I., Churchill, A., Love, J., Axton, R., Moore, T., Clarke, M., Meire, F. and van Heyningen, V. (1999). "Missense mutations in the most ancient residues of the PAX6 paired domain underlie a spectrum of human congenital eye malformations." *Hum Mol Genet* **8**(2): 165-172.

Hanson, I. M. (2001). "Mammalian homologues of the *Drosophila* eye specification genes." *Semin Cell Dev Biol* **12**(6): 475-484.

Hazelett, D. J., Bourouis, M., Walldorf, U. and Treisman, J. E. (1998). "*decapentaplegic* and *wingless* are regulated by *eyes absent* and *eyegone* and interact to direct the pattern of retinal differentiation in the eye disc." *Development* **125**(18): 3741-3751.

Heanue, T. A., Reshef, R., Davis, R. J., Mardon, G., Oliver, G., Tomarev, S., Lassar, A. B. and Tabin, C. J. (1999). "Synergistic regulation of vertebrate muscle development by *Dach2, Eya2*, and *Six1*, homologs of genes required for *Drosophila* eye formation." *Genes Dev* **13**(24): 3231-3243.

Heberlein, U., Singh, C. M., Luk, A. Y. and Donohoe, T. J. (1995). "Growth and differentiation in the *Drosophila* eye coordinated by *hedgehog*." *Nature* **373**(6516): 709-711.

Hill, R. E., Favor, J., Hogan, B. L., Ton, C. C., Saunders, G. F., Hanson, I. M., Prosser, J., Jordan, T., Hastie, N. D. and van Heyningen, V. (1991). "Mouse small eye results from mutations in a *paired*-like homeobox-containing gene." *Nature* **354**(6354): 522-525.

Hoshiyama, D., Suga, H., Iwabe, N., Koyanagi, M., Nikoh, N., Kuma, K., Matsuda, F., Honjo, T. and Miyata, T. (1998). "Sponge Pax cDNA related to *Pax-2/5/8* and ancient gene duplications in the *Pax* family." *J Mol Evol* **47**(6): 640-648.

Ikeda, K., Watanabe, Y., Ohto, H. and Kawakami, K. (2002). "Molecular interaction and synergistic activation of a promoter by Six, Eya, and Dach proteins mediated through CREB binding protein." *Mol Cell Biol* **22**(19): 6759-6766.

Iseki, M., Matsunaga, S., Murakami, A., Ohno, K., Shiga, K., Yoshida, K., Sugai, M., Takahashi, T., Hori, T. and Watanabe, M. (2002). "A blue-light-activated adenylyl cyclase mediates photoavoidance in *Euglena gracilis*." *Nature* **415**(6875): 1047-1051.

Isoldi, M. C., Rollag, M. D., Castrucci, A. M. and Provencio, I. (2005). "Rhabdomeric phototransduction initiated by the vertebrate photopigment melanopsin." *Proc Natl Acad Sci U S A* **102**(4): 1217-1221.

Iwabe, N., Kuma, K. and Miyata, T. (1996). "Evolution of gene families and relationship with organismal evolution: rapid divergence of tissue-specific genes in the early evolution of chordates." *Mol Biol Evol* **13**(3): 483-493.

Jacob, F. (1977). "Evolution and tinkering." *Science* **196**(4295): 1161-1166.

Jarman, A. P. (2000). "Developmental genetics: vertebrates and insects see eye to eye." *Curr Biol* **10**(23): R857-859.

Jaworski, C., Sperbeck, S., Graham, C. and Wistow, G. (1997). "Alternative splicing of *Pax6* in bovine eye and evolutionary conservation of intron sequences." *Biochem Biophys Res Commun* **240**(1): 196-202.

Jean, D., Bernier, G. and Gruss, P. (1999). "*Six6* (*Optx2*) is a novel murine *Six3*-related homeobox gene that demarcates the presumptive pituitary/hypothalamic axis and the ventral optic stalk." *Mech Dev* **84**(1-2): 31-40.

Jun, S., Wallen, R. V., Goriely, A., Kalionis, B. and Desplan, C. (1998). "Lune/eye gone, a Pax-like protein, uses a partial paired domain and a homeodomain for DNA recognition." *Proc Natl Acad Sci U S A* **95**(23): 13720-13725.

Kawakami, K., Ohto, H., Takizawa, T. and Saito, T. (1996). "Identification and expression of *six* family genes in mouse retina." *FEBS Lett* **393**(2-3): 259-263.

Kawakami, K., Sato, S., Ozaki, H. and Ikeda, K. (2000). "*Six* family genes--structure and function as transcription factors and their roles in development." *Bioessays* **22**(7): 616-626.

Kevan, P. G., Chittka, L. and Dyer, A. G. (2001). "Limits to the salience of ultraviolet: lessons from colour vision in bees and birds." *J Exp Biol* **204**(Pt 14): 2571-2580.

Kim, S. S., Zhang, R. G., Braunstein, S. E., Joachimiak, A., Cvekl, A. and Hegde, R. S. (2002). "Structure of the retinal determination protein Dachshund reveals a DNA binding motif." *Structure* **10**(6): 787-795.

Kirby, R. J., Hamilton, G. M., Finnegan, D. J., Johnson, K. J. and Jarman, A. P. (2001). "*Drosophila* homolog of the myotonic dystrophy-associated gene, *SIX5*, is required for muscle and gonad development." *Curr Biol* **11**(13): 1044-1049.

Kloosterman, W. P., Wienholds, E., de Bruijn, E., Kauppinen, S. and Plasterk, R. H. (2006). "*In situ* detection of miRNAs in animal embryos using LNA-modified oligonucleotide probes." *Nat Methods* **3**(1): 27-29.

Kobayashi, M., Toyama, R., Takeda, H., Dawid, I. B. and Kawakami, K. (1998). "Overexpression of the forebrain-specific homeobox gene *six3* induces rostral forebrain enlargement in zebrafish." *Development* **125**(15): 2973-2982.

Kobayashi, M., Nishikawa, K., Suzuki, T. and Yamamoto, M. (2001). "The homeobox protein Six3 interacts with the Groucho corepressor and acts as a transcriptional repressor in eye and forebrain formation." *Dev Biol* **232**(2): 315-326.

Koroma, B. M., Yang, J. M. and Sundin, O. H. (1997). "The *Pax-6* homeobox gene is expressed throughout the corneal and conjunctival epithelia." *Invest Ophthalmol Vis Sci* **38**(1): 108-120.

Kozmik, Z., Kurzbauer, R., Dorfler, P. and Busslinger, M. (1993). "Alternative splicing of *Pax-8* gene transcripts is developmentally regulated and generates isoforms with different transactivation properties." *Mol Cell Biol* **13**(10): 6024-6035.

Kozmik, Z., Czerny, T. and Busslinger, M. (1997). "Alternatively spliced insertions in the paired domain restrict the DNA sequence specificity of Pax6 and Pax8." *Embo J* **16**(22): 6793-6803.

Kozmik, Z., Pfeffer, P., Kralova, J., Paces, J., Paces, V., Kalousova, A. and Cvekl, A. (1999). "Molecular cloning and expression of the human and mouse homologues of the *Drosophila dachshund* gene." *Dev Genes Evol* **209**(9): 537-545.

Krishnan, B., Levine, J. D., Lynch, M. K., Dowse, H. B., Funes, P., Hall, J. C., Hardin, P. E. and Dryer, S. E. (2001). "A new role for cryptochrome in a *Drosophila* circadian oscillator." *Nature* **411**(6835): 313-317.

Kumar, J. P. and Moses, K. (2001a). "Expression of evolutionarily conserved eye specification genes during *Drosophila* embryogenesis." *Dev Genes Evol* **211**(8-9): 406-414.

Kumar, J. P. and Moses, K. (2001b). "EGF receptor and Notch signaling act upstream of Eyeless/Pax6 to control eye specification." *Cell* **104**(5): 687-697.

Kumar, J. P. (2001). Signalling pathways in *Drosophila* and vertebrate retinal development. *Nat Rev Genet* **2**, 846-57.

Laclef, C., Souil, E., Demignon, J. and Maire, P. (2003a). "Thymus, kidney and craniofacial abnormalities in *Six 1*-deficient mice." *Mech Dev* **120**(6): 669-679.

Laclef, C., Hamard, G., Demignon, J., Souil, E., Houbron, C. and Maire, P. (2003b). "Altered myogenesis in *Six1*-deficient mice." *Development* **130**(10): 2239-2252.

Lamey, T. M., Koenders, A. and Ziman, M. (2004). "*Pax* genes in myogenesis: alternate transcripts add complexity." *Histol Histopathol* **19**(4): 1289-1300.

Lebert, M. and Hader, D. P. (1997). "Behavioral mutants of *Euglena gracilis*: functional and spectroscopic characterization." *J Plant Physiol* **151**(2): 188-195.

Li, X., Perissi, V., Liu, F., Rose, D. W. and Rosenfeld, M. G. (2002). "Tissue-specific regulation of retinal and pituitary precursor cell proliferation." *Science* **297**(5584): 1180-1183.

Li, X., Oghi, K. A., Zhang, J., Krones, A., Bush, K. T., Glass, C. K., Nigam, S. K., Aggarwal, A. K., Maas, R., Rose, D. W. and Rosenfeld, M. G. (2003). "Eya protein phosphatase activity regulates Six1-Dach-Eya transcriptional effects in mammalian organogenesis." *Nature* **426**(6964): 247-254.

Loosli, F., Kmita-Cunisse, M. and Gehring, W. J. (1996). "Isolation of a *Pax-6* homolog from the ribbonworm *Lineus sanguineus*." *Proc Natl Acad Sci U S A* **93**(7): 2658-2663.

Loosli, F., Koster, R. W., Carl, M., Krone, A. and Wittbrodt, J. (1998). "*Six3*, a medaka homologue of the *Drosophila* homeobox gene *sine oculis* is expressed in the anterior embryonic shield and the developing eye." *Mech Dev* **74**(1-2): 159-164.

Loosli, F., Winkler, S. and Wittbrodt, J. (1999). "*Six3* overexpression initiates the formation of ectopic retina." *Genes Dev* **13**(6): 649-654.

Lopez-Rios, J., Tessmar, K., Loosli, F., Wittbrodt, J. and Bovolenta, P. (2003). "Six3 and Six6 activity is modulated by members of the groucho family." *Development* **130**(1): 185-195.

Mannini, L., Rossi, L., Deri, P., Gremigni, V., Salvetti, A., Salo, E. and Batistoni, R. (2004). "*Djeyes absent* (*Djeya*) controls prototypic planarian eye regeneration by cooperating with the transcription factor Djsix-1." *Dev Biol* **269**(2): 346-359.

Mardon, G., Solomon, N. M. and Rubin, G. M. (1994). "*dachshund* encodes a nuclear protein required for normal eye and leg development in *Drosophila*." *Development* **120**(12): 3473-3486.

Menaker, M. (2003). "Circadian rhythms. Circadian photoreception." *Science* **299**(5604): 213-214.

Mishra, R., Gorlov, I. P., Chao, L. Y., Singh, S. and Saunders, G. F. (2002). "PAX6, paired domain influences sequence recognition by the homeodomain." *J Biol Chem* **277**(51): 49488-49494.

Neumann, C. J. and Nuesslein-Volhard, C. (2000). "Patterning of the zebrafish retina by a wave of Sonic Hedgehog activity." *Science* **289**(5487): 2137-2139.

Niimi, T., Seimiya, M., Kloter, U., Flister, S. and Gehring, W. J. (1999). "Direct regulatory interaction of the eyeless protein with an eye-specific enhancer in the *sine oculis* gene during eye induction in *Drosophila*." *Development* **126**(10): 2253-2260.

Nilsson, D. E. (1996). "Eye ancestry: old genes for new eyes." *Curr Biol* **6**(1): 39-42.

Nilsson, D. E. (2004). "Eye evolution: a question of genetic promiscuity." *Curr Opin Neurobiol* **14**(4): 407-414.

Oakley, T. H. (2003). "The eye as a replicating and diverging, modular developmental unit." *Trends Ecol. Evol.* **18**: 623-627.

Ohto, H., Kamada, S., Tago, K., Tominaga, S. I., Ozaki, H., Sato, S. and Kawakami, K. (1999). "Cooperation of Six and Eya in activation of their target genes through nuclear translocation of Eya." *Mol Cell Biol* **19**(10): 6815-6824.

Oliver, G., Mailhos, A., Wehr, R., Copeland, N. G., Jenkins, N. A. and Gruss, P. (1995a). "*Six3*, a murine homologue of the *sine oculis* gene, demarcates the most anterior border of the developing neural plate and is expressed during eye development." *Development* **121**(12): 4045-4055.

Oliver, G., Wehr, R., Jenkins, N. A., Copeland, N. G., Cheyette, B. N., Hartenstein, V., Zipursky, S. L. and Gruss, P. (1995b). "Homeobox genes and connective tissue patterning." *Development* **121**(3): 693-705.

Onuma, Y., Takahashi, S., Asashima, M., Kurata, S. and Gehring, W. J. (2002). "Conservation of Pax 6 function and upstream activation by Notch signaling in eye development of frogs and flies." *Proc Natl Acad Sci U S A* **99**(4): 2020-2025.

Ozaki, H., Nakamura, K., Funahashi, J., Ikeda, K., Yamada, G., Tokano, H., Okamura, H. O., Kitamura, K., Muto, S., Kotaki, H., Sudo, K., Horai, R., Iwakura, Y. and Kawakami, K. (2004). "*Six1* controls patterning of the mouse otic vesicle." *Development* **131**(3): 551-562.

Panda, S., Nayak, S. K., Campo, B., Walker, J. R., Hogenesch, J. B. and Jegla, T. (2005). "Illumination of the melanopsin signaling pathway." *Science* **307**(5709): 600-604.

Pauli, T. (2004). Molecular genetic analysis of a *sine oculis* enhancer and the *leventina* gene as a model system to study human macular degeneration in *Drosophila*, Philosophisch-Naturwissenschaftlichen Fakultät der Universität Basel. PhD. dissertation.

Personius, K. E., Nautiyal, J. and Reddy, S. (2005). "Myotonia and muscle contractile properties in mice with *SIX5* deficiency." *Muscle Nerve* **31**(4): 503-505.

Pichaud, F. and Desplan, C. (2002). "*Pax* genes and eye organogenesis." *Curr Opin Genet Dev* **12**(4): 430-434.

Pichaud, F., Treisman, J. and Desplan, C. (2001). "Reinventing a common strategy for patterning the eye." *Cell* **105**(1): 9-12.

Pignoni, F., Hu, B., Zavitz, K. H., Xiao, J., Garrity, P. A. and Zipursky, S. L. (1997). "The eye-specification proteins So and Eya form a complex and regulate multiple steps in *Drosophila* eye development." *Cell* **91**(7): 881-891.

Pineda, D., Gonzalez, J., Callaerts, P., Ikeo, K., Gehring, W. J. and Salo, E. (2000). "Searching for the prototypic eye genetic network: *Sine oculis* is essential for eye regeneration in planarians." *Proc Natl Acad Sci U S A* **97**(9): 4525-4529.

Pineda, D. and Salo, E. (2002a). "Planarian *Gtsix3*, a member of the *Six/so* gene family, is expressed in brain branches but not in eye cells." *Mech Dev* **119 Suppl 1**: S167-171.

Pineda, D., Rossi, L., Batistoni, R., Salvetti, A., Marsal, M., Gremigni, V., Falleni, A., Gonzalez-Linares, J., Deri, P. and Salo, E. (2002b). "The genetic network of prototypic planarian eye regeneration is Pax6 independent." *Development* **129**(6): 1423-1434.

Plachetzki, D. C., Serb, J. M. and Oakley, T. H. (2005). "New insights into the evolutionary history of photoreceptor cells." *Trends Ecol Evol* **20**(9): 465-467.

Punzo, C., Kurata, S. and Gehring, W. J. (2001). "The eyeless homeodomain is dispensable for eye development in *Drosophila*." *Genes Dev* **15**(13): 1716-1723.

Punzo, C., Seimiya, M., Flister, S., Gehring, W. J. and Plaza, S. (2002). "Differential interactions of *eyeless* and *twin of eyeless* with the *sine oculis* enhancer." *Development* **129**(3): 625-634.

Punzo, C., Plaza, S., Seimiya, M., Schnupf, P., Kurata, S., Jaeger, J. and Gehring, W. J. (2004). "Functional divergence between *eyeless* and *twin of eyeless* in *Drosophila* melanogaster." *Development* **131**(16): 3943-3953.

Purcell, P. (2002). Genes involved in early development of the mouse sensory placodes. Thesis, Harvard University.

Quiring, R., Walldorf, U., Kloter, U. and Gehring, W. J. (1994). "Homology of the *eyeless* gene of *Drosophila* to the *Small eye* gene in mice and *Aniridia* in humans." *Science* **265**(5173): 785-789.

Raff, R. A. (1996). The Shape of life: Genes, Development, and the Evolution of Animal Form, University of Chicago Press.

Rayapureddi, J. P., Kattamuri, C., Steinmetz, B. D., Frankfort, B. J., Ostrin, E. J., Mardon, G. and Hegde, R. S. (2003). "Eyes absent represents a class of protein tyrosine phosphatases." *Nature* **426**(6964): 295-298.

Rebay, I., Silver, S. J. and Tootle, T. L. (2005). "New vision from Eyes absent: transcription factors as enzymes." *Trends Genet* **21**(3): 163-171.

Rodrigues, A. B. and Moses, K. (2004). "Growth and specification: fly *Pax6* homologs *eyegone* and *eyeless* have distinct functions." *Bioessays* **26**(6): 600-603.

Ruvkun, G. and Hobert, O. (1998). "The taxonomy of developmental control in *Caenorhabditis elegans*." *Science* **282**(5396): 2033-2041.

Salo, E., Pineda, D., Marsal, M., Gonzalez, J., Gremigni, V. and Batistoni, R. (2002). "Genetic network of the eye in Platyhelminthes: expression and functional analysis of some players during planarian regeneration." *Gene* **287**(1-2): 67-74.

Salvini-Plawen, L. and Mayr, E. (1961). In: Evolutionary biology, vol. 10 (Hecht MK, Steere WC, and Wallace B, eds). New York: Plenum Press; 207-263.

Salvini-Plawen, L. and Mayr, E. (1977). "On the evolution of photoreceptors and eyes." *Evol. Biol.* **10**: 207-263.

Seimiya, M. and Gehring, W. J. (2000). "The *Drosophila* homeobox gene *optix* is capable of inducing ectopic eyes by an *eyeless*-independent mechanism." *Development* **127**(9): 1879-1886.

Seo, H. C., Curtiss, J., Mlodzik, M. and Fjose, A. (1999). "*Six* class homeobox genes in *drosophila* belong to three distinct families and are involved in head development." *Mech Dev* **83**(1-2): 127-139.

Shen, W. and Mardon, G. (1997). "Ectopic eye development in *Drosophila* induced by directed *dachshund* expression." *Development* **124**(1): 45-52.

Silver, S. J., Davies, E. L., Doyon, L. and Rebay, I. (2003). "Functional dissection of *eyes absent* reveals new modes of regulation within the retinal determination gene network." *Mol Cell Biol* **23**(17): 5989-5999.

Silver, S. J. and Rebay, I. (2005). "Signaling circuitries in development: insights from the retinal determination gene network." *Development* **132**(1): 3-13.

Singh, S., Tang, H. K., Lee, J. Y. and Saunders, G. F. (1998). "Truncation mutations in the transactivation region of PAX6 result in dominant-negative mutants." *J Biol Chem* **273**(34): 21531-21541.

Singh, S., Chao, L. Y., Mishra, R., Davies, J. and Saunders, G. F. (2001). "Missense mutation at the C-terminus of PAX6 negatively modulates homeodomain function." *Hum Mol Genet* **10**(9): 911-918.

Singh, S., Mishra, R., Arango, N. A., Deng, J. M., Behringer, R. R. and Saunders, G. F. (2002). "Iris hypoplasia in mice that lack the alternatively spliced *Pax6(5a)* isoform." *Proc Natl Acad Sci U S A* **99**(10): 6812-6815.

Stanescu, D., Iseli, H. P., Schwerdtfeger, K., Ittner, L. M., Reme, C. E. and Hafezi, F. (2007). "Continuous expression of the homeobox gene *Pax6* in the ageing human retina." *Eye* **21**: 90-93.

Stierwald, M., Yanze, N., Bamert, R. P., Kammermeier, L. and Schmid, V. (2004). "The *Sine oculis/Six* class family of homeobox genes in jellyfish with and without eyes: development and eye regeneration." *Dev Biol* **274**(1): 70-81.

Sun, H., Rodin, A., Zhou, Y., Dickinson, D. P., Harper, D. E., Hewett-Emmett, D. and Li, W. H. (1997). "Evolution of paired domains: isolation and sequencing of jellyfish and hydra *Pax* genes related to *Pax-5* and *Pax-6*." *Proc Natl Acad Sci U S A* **94**(10): 5156-5161.

Takechi, M. and Kawamura, S. (2005). "Temporal and spatial changes in the expression pattern of multiple red and green subtype *opsin* genes during zebrafish development." *J Exp Biol* **208**(Pt 7): 1337-1345.

Tang, H. K., Singh, S. and Saunders, G. F. (1998). "Dissection of the transactivation function of the transcription factor encoded by the eye developmental gene *PAX6*." *J Biol Chem* **273**(13): 7210-7221.

Tarpin, M., Gehring, W. J. and Bierne, J. (1999). "Reverse homeosis in homeotically reconstructed ribbonworms." *Proc Natl Acad Sci U S A* **96**(21): 11900-11903.

Tarpin, M., Gehring, W. J. and Bierne, J. (2002). "Conversion of a postocellar into an ocellar region as a transdetermination event occurring in adult ribbonworms." *Mech Dev* **118**(1-2): 39-44.

Tavsanli, B. C., Ostrin, E. J., Burgess, H. K., Middlebrooks, B. W., Pham, T. A. and Mardon, G. (2004). "Structure-function analysis of the *Drosophila* retinal determination protein Dachshund." *Dev Biol* **272**(1): 231-247.

Terakita, A. (2005). "The opsins." *Genome Biol* **6**(3): 213.

Tessmar-Raible, K. (2004). The evolution of sensory and neurosecretory cell types in bilaterian brains, Philipps-Universität Marburg. PhD dissertation.

Thomsen, R., Nielsen, P. S. and Jensen, T. H. (2005). "Dramatically improved RNA *in situ* hybridization signals using LNA-modified probes." *Rna* **11**(11): 1745-1748.

Tjian, R. and Maniatis, T. (1994). "Transcriptional activation: a complex puzzle with few easy pieces." *Cell* **77**(1): 5-8.

Tomarev, S. I., Callaerts, P., Kos, L., Zinovieva, R., Halder, G., Gehring, W. and Piatigorsky, J. (1997). "Squid *Pax-6* and eye development." *Proc Natl Acad Sci U S A* **94**(6): 2421-2426.

Ton, C. C., Hirvonen, H., Miwa, H., Weil, M. M., Monaghan, P., Jordan, T., van Heyningen, V., Hastie, N. D., Meijers-Heijboer, H., Drechsler, M. and et al. (1991). "Positional cloning and characterization of a paired box- and homeobox-containing gene from the *aniridia* region." *Cell* **67**(6): 1059-1074.

Tootle, T. L., Silver, S. J., Davies, E. L., Newman, V., Latek, R. R., Mills, I. A., Selengut, J. D., Parlikar, B. E. and Rebay, I. (2003). "The transcription factor Eyes absent is a protein tyrosine phosphatase." *Nature* **426**(6964): 299-302.

Toy, J., Yang, J. M., Leppert, G. S. and Sundin, O. H. (1998). "The *optx2* homeobox gene is expressed in early precursors of the eye and activates retina-specific genes." *Proc Natl Acad Sci U S A* **95**(18): 10643-10648.

Toy, J. and Sundin, O. H. (1999). "Expression of the *optx2* homeobox gene during mouse development." *Mech Dev* **83**(1-2): 183-186.

Treisman, J., Harris, E. and Desplan, C. (1991). "The paired box encodes a second DNA-binding domain in the paired homeodomain protein." *Genes Dev* **5**(4): 594-604.

Van der Horst, G. T., Muijtjens, M., Kobayashi, K., Takano, R., Kanno, S., Takao, M., de Wit, J., Verkerk, A., Eker, A. P., van Leenen, D., Buijs, R., Bootsma, D., Hoeijmakers, J. H. and Yasui, A. (1999). "Mammalian Cry1 and Cry2 are essential for maintenance of circadian rhythms." *Nature* **398**(6728): 627-630.

Vernet, G. (1970). "Ultrastructure des photorecepteurs de *Lineus ruber* (O. F. Müller) (Heteronemertes Lineidae)." *Z. zellforsch. Mikrosk. Anat.* **104**: 494-506.

Vernet, G. (1974). "Etude ultrastructurale des cellules présumées photoréceptrices dans les ganglions cérébroides des *Lineidae* (Hétéronemertes)." *Ann. Sci. Nat. Zool.* **16**: 27-36.

Walther, C. and Gruss, P. (1991). "*Pax-6*, a murine paired box gene, is expressed in the developing CNS." *Development* **113**(4): 1435-1449.

Wargelius, A., Seo, H. C., Austbo, L. and Fjose, A. (2003). "Retinal expression of zebrafish *six3.1* and its regulation by *Pax6.*" *Biochem Biophys Res Commun* **309**(2): 475-481.

Wawersik, S. and Maas, R. L. (2000). "Vertebrate eye development as modeled in *Drosophila*." *Hum Mol Genet* **9**(6): 917-925.

Wilson, D., Sheng, G., Lecuit, T., Dostatni, N. and Desplan, C. (1993). "Cooperative dimerization of paired class homeo domains on DNA." *Genes Dev* **7**(11): 2120-2134.

Xu, W., Rould, M. A., Jun, S., Desplan, C. and Pabo, C. O. (1995). "Crystal structure of a paired domain-DNA complex at 2.5 A resolution reveals structural basis for Pax developmental mutations." *Cell* **80**(4): 639-650.

Xu, P. X., Cheng, J., Epstein, J. A. and Maas, R. L. (1997). "Mouse *Eya* genes are expressed during limb tendon development and encode a transcriptional activation function." *Proc Natl Acad Sci U S A* **94**(22): 11974-11979.

Yanowitz, J. L., Shakir, M. A., Hedgecock, E., Hutter, H., Fire, A. Z. and Lundquist, E. A. (2004). "*UNC-39*, the *C. elegans* homolog of the human myotonic dystrophy-associated homeodomain protein Six5, regulates cell motility and differentiation." *Dev Biol* **272**(2): 389-402.

Zheng, W., Huang, L., Wei, Z. B., Silvius, D., Tang, B. and Xu, P. X. (2003). "The role of *Six1* in mammalian auditory system development." *Development* **130**(17): 3989-4000.

Zuber, M. E., Perron, M., Philpott, A., Bang, A. and Harris, W. A. (1999). "Giant eyes in *Xenopus laevis* by overexpression of *XOptx2*." *Cell* **98**(3): 341-352.

Zuber, M. E., Gestri, G., Viczian, A. S., Barsacchi, G. and Harris, W. A. (2003). "Specification of the vertebrate eye by a network of eye field transcription factors." *Development* **130**(21): 5155-5167.

IV. Materials and Methods

Standard molecular methods, such as PCR product purification, alkaline phosphatase treatment, ligation, preparation and transformation of competent cells, plasmid miniprep, DNA digestion, phenol-chloroform extraction of DNA were performed according to the "Maniatis" (Sambrook and Russel, 2001) and to the manufacturer protocols. They will not be further described.

1. mRNA extraction

The mRNAs from various tissues were extracted using the Dynabeads® mRNA DIRECT KIT™ from DYNAL® Biotech. This kit allows a reliable and rapid isolation of mRNAs directly from tissues. The isolation protocol, performed in a single tube, is based on the ability of magnetic particles that are coupled with oligo(dT)25 to specifically retain poly-A mRNAs. Highly purified poly-A mRNAs are isolated and are ready-to-go for further downstream analysis or reverse transcription step.

2. Reverse transcription

cDNA were synthesized with the SuperScriptTM ^{III} First-Strand Synthesis System for RT-PCR from Invitrogen. When the cDNA were synthesized for subsequent real-time PCR, we used random hexamers as primers for the reverse transcription, whereas we used poly-dT primers when the cDNA were synthesized for subsequent PCR experiments with specific- or degenerated- primers.

3. Degenerated PCR

First step was always a denaturing step for 2 minutes at 94°C, followed by 30 to 50 cycles with the pimer-specific annealing temperature for 30 seconds and the required elongation temperature for 1 to several minutes, depending on the expected size of the PCR product (1 minute for approximatively 1kb). We mostly used the Taq polymerase from Roche. The program was finished by a final elongation step at 72°C resulting in the following scheme:

initial denaturation	2min	at	94°C		
denaturation	30sec	at	94°C	}	30-50X
annealing	30sec	at	X°C	}	
elongation	Ysec	at	72°C	}	
final elongation	5min	at	72°C		
stop	∞	at	8°C		

PCR reactions were mostly done in a nested manner, using 1µl of the 1st round PCR reaction in a final volume of 20 µl. PCR products were subsequently purified by agarose gel electrophoresis. Bands of interest were cut out of the gel. The DNA was eluted from the gel by using the QIAEX gel extraction kit.

4. Cloning of PCR products

The purified PCR products were cloned in the pCR®II-TOPO® vector from the TOPO TA cloning ® kit from Invitrogen. This vector kit makes use of the single 3' A-overhang, which is added by the Taq polymerase onto the PCR amplified products. PCR products with these 3' A-overhangs can easily be cloned into the pCR®II-TOPO® vector, which contains a single 3' T-overhang. Hence, this kit provides a highly efficient, one-step strategy cloning for the direct insertion of a PCR product into a plasmid. As M13 Reverse and M13 Forward priming sites are surrounding the multiple cloning site of the pCR®II-TOPO® vector, we used M13 Reverse and M13 Forward primers to sequence our insert.

5. DNA sequencing

We used the BigDye® Terminator reaction kit, PE Applied Biosystems to labeled DNA fragments we wanted to sequence. The PCR reactions were done according to the manufacturer. The sequencing analysis was done on an ABI PRISM 310 Genetic Analyzer.

6. 3' and 5' RACE PCR

The full coding sequences were obtained by RACE (rapid amplification of cDNA ends) on cDNA prepared from polyadenylated RNA extracted from various adult and regenerating *L.sanguineus* tissues. We used the SMART RACE cDNA amplification kit from Clontech.

7. Gene expression level analysis by real-time PCR

The gene expression level analysis by real-time PCR was done on a Light-Cycler II instrument from Roche. Tissue-specific cDNA were used. Tissue-specific values of the target gene were normailized against elongation factor expression levels. The primers used were designed with a specific program available from the website www.genscript.com/ssl-bin/app/primer. The Light Cycler FastStart DNA Master SYBR Green I mix from Roche was used for the real-time PCR reactions. The standars amplification protocol was: 15 minutes at 95°C, 15 seconds at 94°C, 20 seconds at 58°C and 10 seconds at 72°C, over 40 cycles. The fluorescence of the amplified products was detected 5 seconds at 76°C, after the elongation step.

specific primers used for the real-time PCR experiments:

qRT.EF1 alpha.Sense	GGCAAACCTTCCGAGTGGCGGGTAATCG
qRT.EF1 alpha.Asense	CCGGGTGATTGAGAATGATGACCTGTGC
qRT.OTX.Sense	GAGCACCCTACTCCGTCAAC
qRT.OTX.Asense	CGACTCTAGCGTGTACCAACC
qRT.CDX.Sense	TGACTGGACCACCATCGTTA
qRT.CDX.Asense	ATGTTTGGGCAAGGCTGT

8. Preparation of RNA probes

In vitro transcription:

200-400 ng of DNA (purified PCR product with T7 promoter at the 3'end)
2 µL of 10X transcription buffer
2 µL of DIG-NTP labeling Mix from Boehringer
0.5 µL RNase inhibitor
2 µL of T7 RNA polymerase
add dH2O to 20 µL

incubate 2-3 hours at 37°C

Remove DNA template by adding 1 µL of RNAse free DNAse I and incubate for another 15 min at 37°C.

Stop the reaction by adding 2 µL 0,2 M EDTA pH 8.

Adjust volume to 50 µL by adding DEPC treated H2O.

Precipitate by adding 25 µL 7.8 M NH_4Ac and 150 µL 100% EtOH.

Incubate 30 min at -20°C

Centrifuge 15 min at 4°C

Wash with ice cold 80% EtOH

Resuspend visible pellet in 25 µL DEPC treated H2O, analyze 1 µL on agarose gel

Add 76 µL Hibridization Buffer (HB) and store at -20°C

Usually 4 µL of this probe/HB is used in 500 µL HB.

9. *in situ* hybridization protocol

Tissue Preparation

1. Rinse with PBS 2X
2. Fixation in 4% PFA in PBS, 5 min
3. Rinse with PBS 2X, wash 2X 2 min in PBS
4. Cystein chloride treatment 0,1M in PBS: rinse once, and then incubate 10 min at RT
5. Rinse with PBS 2X, wash 3X 2 min in PBS
6. Fixation in 4% PFA in PBS, 20 min
7. Rinse with PBS 2X, wash 1X 2 min in PBS
8. Wash 3X 2 min in PBT
9. Rinse 3X in methanol

From now, the samples can be stored at -20°C

In situ procedures

1. Rehydrate: 2 min with 75% methanol in PBT/ 2 min with 50% methanol in PBT/ 2 min with 25% methanol in PBT
2. Wash 3X in PBT
3. Post-fixation in 4% PFA in PBT, 10 min
4. Rinse with PBT 2X, wash 4X 2 min in PBT
5. Proteinase K treatment: 10µg.ml^{-1} in PBT, 10 min
6. Rinse once with PBT
7. Post-fixation in 4% PFA and 0,2% glutaraldehyde in PBT, 10 min
8. Rinse 2X with PBT, wash 3X 2min in PBT
9. Rinse 2X with RIPA buffer, wash 3X 10 in RIPA
10. Rinse 2X with PBT, wash 2X 2 min in PBT

11. Wash in 50% HB/ 50% PBT for 2X 5min
12. Wash 2X in HB
13. Pre-hybridize in HB at 56°C for 2-3 hours
14. Replace Pre-hybridization solution with HB that contains the RNA probe, incubate over night at 56°C
15. Wash in HB at 56°C, 5 min
16. Wash in WS1 2X 20 min at 56°C
17. Wash in WS2 2X 10 min at 56°C
18. Wash in WS3 2X 20 min at 56°C
19. Rinse with PBT at RT
20. Wash 3X 5 min with PBT
21. Rinse with blocking solution at RT
22. Incubate 90 min in blocking solution at RT
23. Incubate for 2 hours maximum at RT with pre-adsorbed antibody 1:2000 in blocking solution or over night at 4°C with pre-adsorbed antibody 1:5000
24. Rinse 2X with MAB
25. Wash 3X 10 min in MAB
26. Rinse 2X with AP 9.5 buffer
27. Wash 2X 5 min in AP 9.5 buffer
28. Color detection: 4,5 µL NBT and 3,5 µL BCIP in 1 mL AP 9.5 buffer
29. Monitor the staining. When desired, stop it by several rinse in PBT
30. Mount in Glycergel

Solutions:

PBS: 130 mM NaCl, 7 mM Na_2HPO_4, 3mM NaH_2PO_4

4% PFA: use 16% Paraformaldehyde, EM grade.

PBT: 0,1 % Triton in PBSl

RIPA buffer: for 20 mL: 14 mL dH20, 1 mL NaCl 5M, 2 mL NP40 10%, 2 mL DOC 5%, 660 µL TrisCl pH8 1M, 300 µL SDS 20%, 40 µL EDTA 0,5M.

HB: for 50 mL: 25 mL formamide, 12,5 mL SSC 20X, 11,5 mL dH2O, 500 µL Tween 10%, 500 µL heparine 10 mg/mL, 5 µL salmon sperm DNA and 5 mg yeast (or E.coli) RNA.

WS1: 50% formamide, 2X SSCT. For 12 mL: 6 mL formamide, 1,2 mL SSC 20X, 120 µL Tween 10%, 4,68 mL dH20

WS2: 2X SSCT. For 12 mL: 1,2 mL SSC, 120 µL Tween 10%, 10,68 mL dH20

WS3: 0.2 SSCT. For 12 mL: 120 µL SSC, 120 µLTween 10%, 11,76 mL dH20

Blocking solution: 1% blocking reagent from Boehringer in MAB

MAB: 100mM maleic acid pH7.5, 150 mM NaCl. For 200 mL: 20 mL maleic acid 1M pH7.5, 6 mL NaCl 5M, 2 mL Tween 10%, 172 mL dH20.

A.P 9.5 buffer: 100 mM NaCl, 50 mM MgCl2, 100 mM TrisCl pH 9.5, 0,1% Tween, 1 mM levamisol. For 10 mL: 200 µL NaCl 5M, 500 µL MgCl2 1M, 1 mL TrisCl pH 9.5 1M, 50 µL Tween 20%, 10 µL levamisol 1M in dH20.

Stock: NBT: 75 mg/mL in 70% DMF

BCIP: 50 mg/mL

10. Antibody staining

Same tissue preparation than the one from the *in situ* hybridization protocol, followed by:

Rehydrate: 2 min with 75% methanol in PBT/ 2 min with 50% methanol in PBT/ 2 min with 25% methanol in PBT

Wash 3X in PBT

Post-fixation in 5% formaldehyde in PBT, 20 min

Rinse with PBT 2X, wash 4X 2 min in PBT

Rinse twice and incubate 30 min in PBT containing 0,1% Triton, 0,1% Tween20, 0,1% NP40, 0,1% DOC

Rinse once with PBT

Rinse with blocking solution

Incubate 90 min in blocking solution

Incubate over night at 4°C with the first antibody at the appropriate dilution

Wash 5X 10 min in MAB

Incubate 90 min in blocking solution

Incubate 2 hours at RT with the second antibody at the appropriate dilution

Wash 5X 10 min in MAB

For color detection, when the second Ab was conjugated with AP:

Rinse 2X with AP 9.5 buffer

Wash 2X 5 min in AP 9.5 buffer

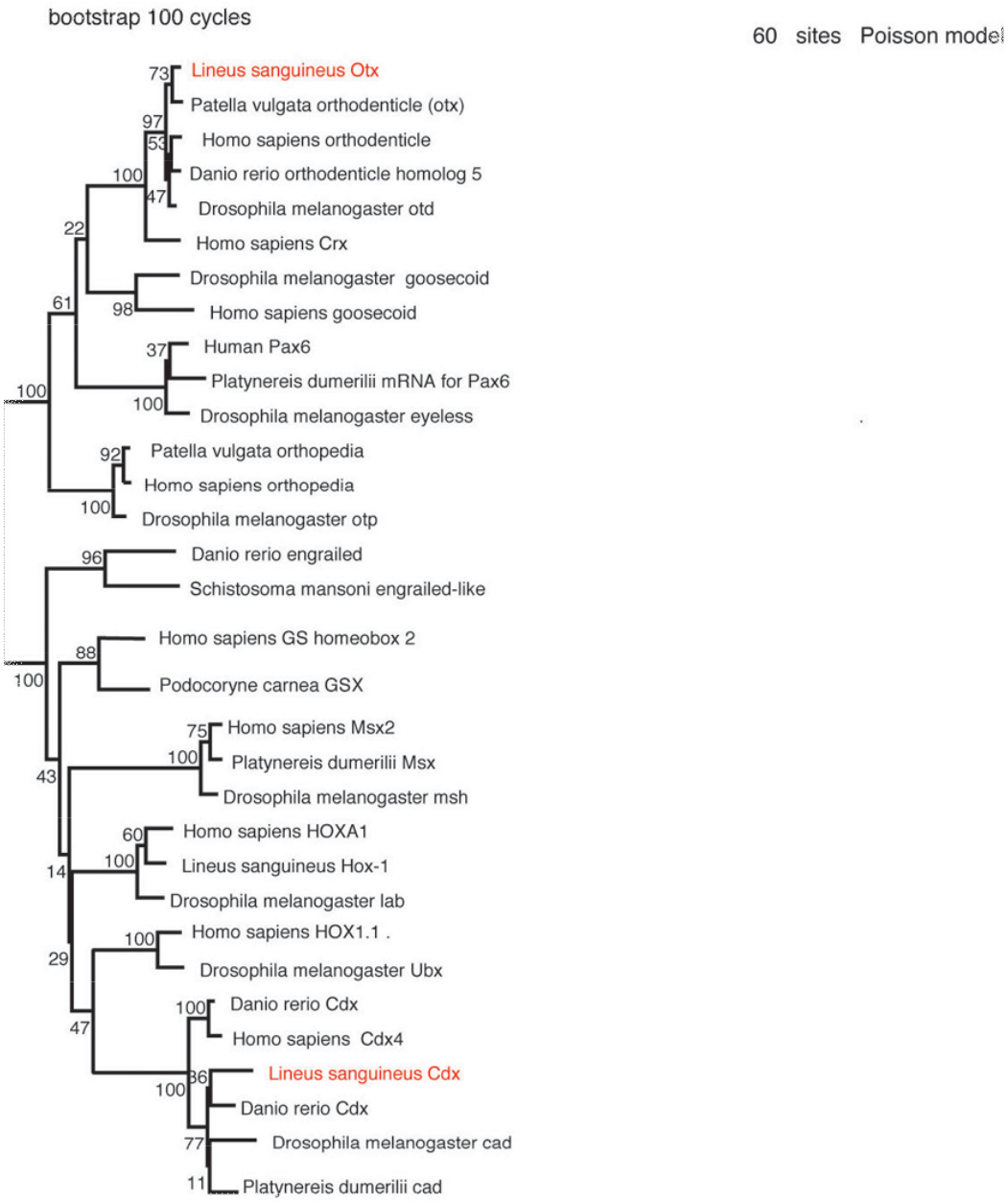
Color detection: 4,5 μL NBT and 3,5 μL BCIP in 1 mL AP 9.5 buffer
Monitor the staining. When desired, stop it by several rinse in PBT
Mount in Glycergel

11. Experimental animals

L.sanguineus and *L.viridis* were collected in Roscoff, France, while *L.lacteus* were collected in Banyuls, France. They were usually kept in the dark at 14°C. The glass bottles, where they were kept, were filled with sea water that was collected from Roscoff. It is known that "artificial sea water" made with salt mix is harmful for the worms.

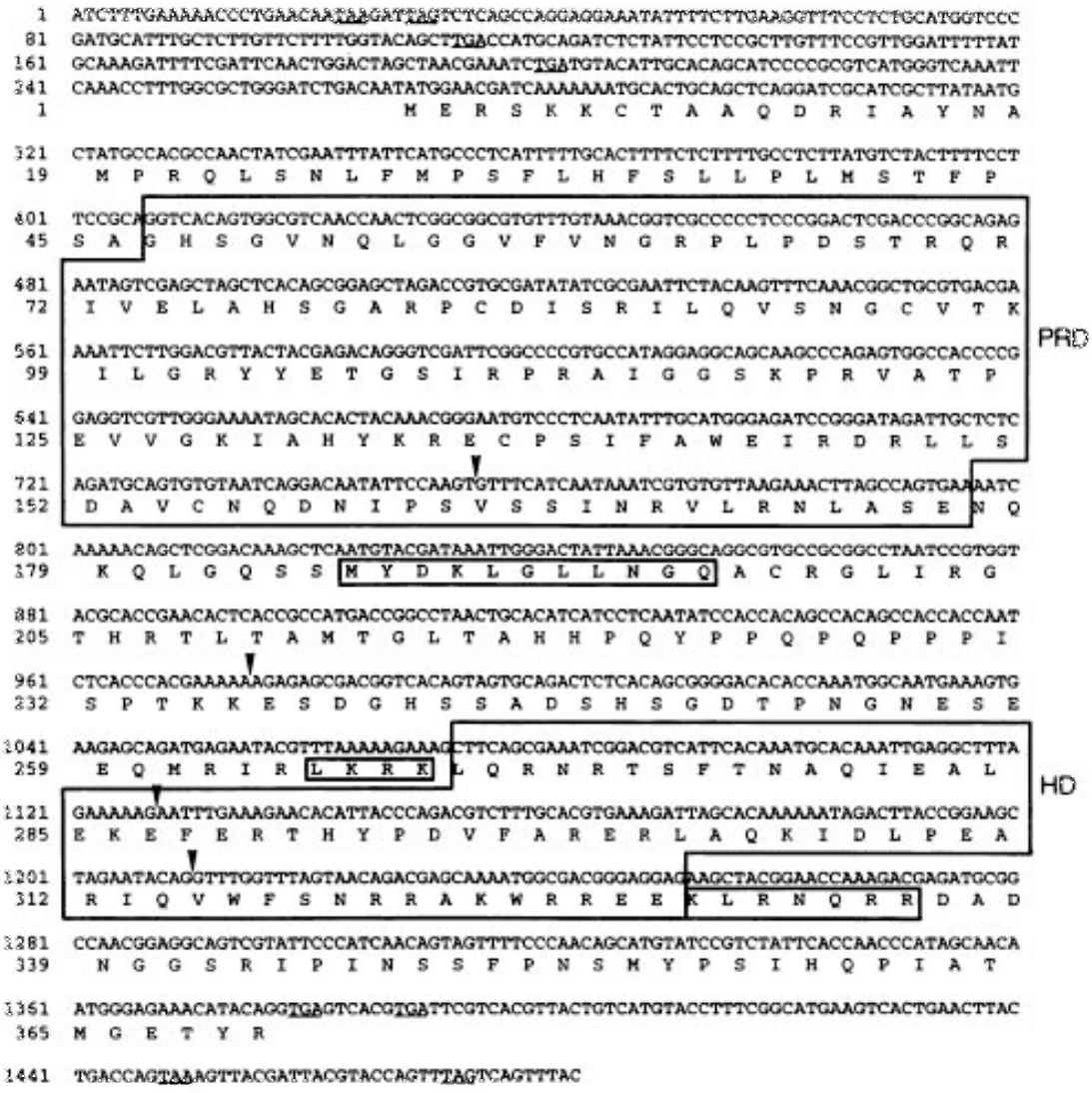
Sambrook, and Rusell. (2001). Molecular cloning. A Laboratory Manual, 3rd edition. (New York: Cold Spring Harbor Laboratory Press).

V. Appendix



Appendix 1. Phylogenetic tree showing *Ls-Otx* and *Ls-Cdx* genes in relation to other homeobox-containing genes

The phylogenetic tree was constructed by using the neighbour-joining (NJ) method. All the amino acids from the HD were used. Poisson-correction distance was used to draw the NJ tree. The *Ls-Otx* and *Ls-Cdx* genes are highlighted in red.



Appendix 2. The *Ls-Pax-6* sequence, which was available before this PhD work, and the deduced amino acid sequence of its putative ORF.

The paired domain (PRD) and the homeodomain (HD) are indicated with boxes. The conserved motif in the linker region and the conserved amino acids flanking the homeodomain are framed in. The splice sites are indicated by arrowheads. On the basis of sequence homology there may be an additional splice site in the first codon of the paired box. In-frame stop codons are underlined.

Ls-Pax6 ACAGCTTGACCATGCAGATCTCTATTCCTCCGCTTGTTTCCGTTGGATTTTTATGCAAAG 60
Ls-Pax6' ACAGCTTGACCATGCAGATCTCTATTCCTCCGCTTGTTTCCGTTGGATTTTTATGCAAAG 60

Ls-Pax6 ATTTTCGATTCAACTGGACTAGCTAACGAAATCTGATGTACATTGCACAGCATCCCCGCG 120
Ls-Pax6' ATTTTCGATTCAACTGGACTAGCTAACGAAATCTGATGTACATTGCACAGCATCCCCGCG 120

Ls-Pax6 TCATGGGTCAAATTCAAACCTTTGGCGCTGGGATCTGACAAT**ATG**GAACGATCAAAAAAA 180
Ls-Pax6' TCATGGGTCAAATTCAAACCTTTGGCGCTGGGATCTGACAAT**ATG**GAACGATCAAAAAAA 180

Ls-Pax6 TGCACTGCAGCTCAGGATCGCATCGCTTATAATGCTATGCCACGCCAACTATCGAATTTA 240
Ls-Pax6' TGCACTGCAGCTCAGGATCGCATCGCTTATAATGCTATGCCACGCCAACTATCGAATTTA 240

Ls-Pax6 TTCATGCCCTCATTTTTGCACTTTTCTCTTTTGCCTCTTATGTCTACTTTTCCTTCCGCA 300
Ls-Pax6' TTCATGCCCTCATTTTTGCACTTTTCTCTTTTGCCTCTTATGTCTACTTTTCCTTCCGCA 300

Ls-Pax6 GGTCACAGTGGCGTCAACCAACTCGGCGGCGTGTTTGTAAACGGTCGCCCCCTCCCGGAC 360
Ls-Pax6' GGTCACAGTGGCGTCAACCAACTCGGCGGCGTGTTTGTAAACGGTCGCCCCCTCCCGGAC 360

Ls-Pax6 TCGACCCGGCAGAGAATAGTCGAGCTAGCTCACAGCGGAGCTAGACCGTGCGATATATCG 420
Ls-Pax6' TCGACCCGGCAGAGAATAGTCGAGCTAGCTCACAGCGGAGCTAGACCGTGCGATATATCG 420

Ls-Pax6 CGAATTCTACAAGTTTCAAACGGCTGCGTGAGCAAAATTCTTGGACGTTACTACGAGACA 480
Ls-Pax6' CGAATTCTACAAGTTTCAAACGGCTGCGTGAGCAAAATTCTTGGACGTTACTACGAGACA 480

Ls-Pax6 GGGTCGATTCGGCCCCGTGCCATAGGAGGCAGCAAGCCCAGAGTGGCCACCCCGGAGGTC 540
Ls-Pax6' GGGTCGATTCGGCCCCGTGCCATAGGAGGCAGCAAGCCCAGAGTGGCCACCCCGGAGGTC 540

Ls-Pax6 GTTGGGAAAATAGCACACTACAAACGGGAATGTCCCTCAATATTTGCATGGGAGATCCGG 600
Ls-Pax6' GTTGGGAAAATAGCACACTACAAACGGGAATGTCCCTCAATATTTGCATGGGAGATCCGG 600

Ls-Pax6 GATAGATTGCTCTCAGATGCAGTGTGTAATCAGGACAATATTCCAAGTGTTTCATCAATA 660
Ls-Pax6' GATAGATTGCTCTCAGATGCAGTGTGTAATCAGGACAATATTCCAAGTGTTTCATCAATA 660

Ls-Pax6 AATCGTGTGTTAAGAAACTTAGCCAGTGAAAATCAAAAACAGCTCGGACAAAGCTCAATG 720
Ls-Pax6' AATCGTGTGTTAAGAAACTTAGCCAGTGAAAATCAAAAACAGCTCGGACAAAGCTCAATG 720

Ls-Pax6 TACGATAAATTGGGACTATTAAACGGGCAGGCGTGGCCGCGGCCTAATCCGTGGTACGCA 780
Ls-Pax6' TACGATAAATTGGGACTATTAAACGGGCAGGCGTGGCCGCGGCCTAATCCGTGGTACGCA 780

Ls-Pax6 CCGAACACTCACCCGGCCATGACCGGCCTAACTGCACATCATCCTCAATATCCACCACAG 840
Ls-Pax6' CCGAACACTCACCCGGCCATGACCGGCCTAACTGCACATCATCCTCAATATCCACCACAG 840

Ls-Pax6 CCACAGCCACCACCAATCTCACCCACGAAAAAAGAGAGCGACGGTCACAGTAGTGCAGAC 900
Ls-Pax6' CCACAGCCACCACCAATCTCACCCACGAAAAAAGAGAGCGACGGTCACAGTAGTGCAGAC 900

Ls-Pax6 TCTCACAGCGGGGACACACCAAATGGCAATGAAAGTGAAGAGCAGATGAGAATACGTTTA 960
Ls-Pax6' TCTCACAGCGGGGACACACCAAATGGCAATGAAAGTGAAGAGCAGATGAGAATACGTTTA 960

Ls-Pax6 AAAAGAAAGCTTCAGCGAAATCGGACGTCATTCACAAATGCACAAATTGAGGCTTTAGAA 1020
Ls-Pax6' AAAAGAAAGCTTCAGCGAAATCGGACGTCATTCACAAATGCACAAATTGAGGCTTTAGAA 1020

Ls-Pax6 AAAGAATTTGAAAGAACACATTACCCAGACGTCTTTGCACGTGAAAGATTAGCACAAAAA 1080
Ls-Pax6' AAAGAATTTGAAAGAACACATTACCCAGACGTCTTTGCACGTGAAAGATTAGCACAAAAA 1080

Ls-Pax6 ATAGACTTACCGGAAGCTAGAATACAGGTTTGGTTTAGTAACAGACGAGCAAAATGGCGA 1140
Ls-Pax6' ATAGACTTACCGGAAGCTAGAATACAGGTTTGGTTTAGTAACAGACGAGCAAAATGGCGA 1140

Ls-Pax6 CGGGAGGAGAAGCTACGGAACCAAAGACGAGATGCGGCCAACGGAGGCAGTCGTATTCCC 1200
Ls-Pax6' CGGGAGGAGAAGCTACGGAACCAAAGACGAGATGCGGCCAACGGAGGCAGTCGTATTCCC 1200

Ls-Pax6 ATCAACAGTAGTTTTCCCAACAGCATGTATCCGTCTATTCACCAACCCATAGCAACAATG 1260
Ls-Pax6' ATCAACAGTAGTTTTCCCAACAGCATGTATCCGTCTATTCACCAACCCATAGCAACAATG 1260

Ls-Pax6 GGAGAAACATACAGCATGGCCCCAGTGGCAAATTATAGTCTGTCCAATAGCATCCCTCCC 1320
Ls-Pax6' GGAGAAACATACAGCATGGCCCCAGTGGCAAATTATAGTCTGTCCAATAGCATCCCTCCC 1320

Ls-Pax6 AACCCAGCTTGTCTACAGTCGACGAATTCACCATCATCATATTCATGTATGTTACCAGGA 1380
Ls-Pax6' AACCCAGCTTGTCTACAGTCGACGAATTCACCATCATCATATTCATGTATGTTACCAGGA 1380

Ls-Pax6 GGATATACAGGAACAGCTAGAAGCTATGACCCCCTGAGCTTGAGTAGTTACTCCCGACCT 1440
Ls-Pax6' GGATATACAGGAACAGCTAGAAGCTATGACCCCCTGAGCTTGAGTAGTTACTCCCGACCT 1440

Ls-Pax6 ACCTGTAACCCCCACGCAGCAGCAAGCATGCAGAGTCACATGACGCATCAAGCAAATGGC 1500
Ls-Pax6' ACCTGTAACCCCCACGCAGCAGCAAGCATGCAGAGTCACATGACGCATCAAGCAAATGGC 1500

```
Ls-Pax6    GCTTCAACCGG------------------------------CTTAATATCGCCGGGCGTCTCCGTACCAGTACAAG 1546
Ls-Pax6'   GCTTCAACCGGTATGATGGCCTGGGCTTAATATCGCCGGGCGTCTCCGTACCAGTACAAG 1560

Ls-Pax6    TCCCAGGAGGCGGATCAGCTCAGGACGTGGCCCAAGCACACATGGCCTCTCATATGGCCT 1606
Ls-Pax6'   TCCCAGGAGGCGGATCAGCTCAGGACGTGGCCCAAGCACACATGGCCTCTCATATGGCCT 1620

Ls-Pax6    CACAGTATTGGTCAAGGATACAGTGACCTTTGACCATGTTTGGTGACCTTGAACATTGAA 1666
Ls-Pax6'   CACAGTATTGGTCAAGGATACAGTGACCTTTGACCATGTTTGGTGACCTTGAACATTGAA 1680

Ls-Pax6    AGCCCCGGATGAAGCGAAAAGGCATCATTGGGTGAAGTTTAAGATAAACTCTTATTGTGC 1726
Ls-Pax6'   AGCCCCGGATGAAGCGAAAAGGCATCATTGGGTGAAGTTTAAGATAAACTCTTATTGTGC 1740

Ls-Pax6    AATTGGCATGGAAA 1740
Ls-Pax6'   AATTGGCATGGAAA 1754
```

Appendix 3. Comparison of the *Ls-Pax6* splice variant sequence with the canonical one.

The sequence of the *LsPax-6* splice mRNA variant (referred here as *Ls-Pax6'*), which has been identified during this PhD work, differs from the canonical *Ls-Pax6* (referred here as *Ls-Pax6*) by an insertion of 14 nucleotides only. This insertion is responsible for a frame shift, which leads to the creation of an earlier stop. The putative start codons are highlighted in blue. The putative stop codons are highlighted in red.

```
PAX6.1          MERSKKCTAAQDRIAYNAMPRQLSNLFMPSFLHFSLLPLMSTFPSAGHSGVNQLGGVFVN 60
PAX6.2          MERSKKCTAAQDRIAYNAMPRQLSNLFMPSFLHFSLLPLMSTFPSAGHSGVNQLGGVFVN 60
                ************************************************************

PAX6.1          GRPLPDSTRQRIVELAHSGARPCDISRILQVSNGCVSKILGRYYETGSIRPRAIGGSKPR 120
PAX6.2          GRPLPDSTRQRIVELAHSGARPCDISRILQVSNGCVSKILGRYYETGSIRPRAIGGSKPR 120
                ************************************************************

PAX6.1          VATPEVVGKIAHYKRECPSIFAWEIRDRLLSDAVCNQDNIPSVSSINRVLRNLASENQKQ 180
PAX6.2          VATPEVVGKIAHYKRECPSIFAWEIRDRLLSDAVCNQDNIPSVSSINRVLRNLASENQKQ 180
                ************************************************************

PAX6.1          LGQSSMYDKLGLLNGQAWPRPNPWYAPNTHPAMTGLTAHHPQYPPQPQPPPISPTKKESD 240
PAX6.2          LGQSSMYDKLGLLNGQAWPRPNPWYAPNTHPAMTGLTAHHPQYPPQPQPPPISPTKKESD 240
                ************************************************************

PAX6.1          GHSSADSHSGDTPNGNESEEQMRIRLKRKLQRNRTSFTNAQIEALEKEFERTHYPDVFAR 300
PAX6.2          GHSSADSHSGDTPNGNESEEQMRIRLKRKLQRNRTSFTNAQIEALEKEFERTHYPDVFAR 300
                ************************************************************

PAX6.1          ERLAQKIDLPEARIQVWFSNRRAKWRREEKLRNQRRDAANGGSRIPINSSFPNSMYPSIH 360
PAX6.2          ERLAQKIDLPEARIQVWFSNRRAKWRREEKLRNQRRDAANGGSRIPINSSFPNSMYPSIH 360
                ************************************************************

PAX6.1          QPIATMGETYSMAPVANYSLSNSIPPNPACLQSTNSPSSYSCMLPGGYTGTARSYDPLSL 420
PAX6.2          QPIATMGETYSMAPVANYSLSNSIPPNPACLQSTNSPSSYSCMLPGGYTGTARSYDPLSL 420
                ************************************************************

PAX6.1          SSYSRPTCNPHAAASMQSHMTHQANGASTGLISPGVSVPVQVPGGGSAQDVAQAHMASHM 480
PAX6.2          SSYSRPTCNPHAAASMQSHMTHQANGASTGMMAWA------------------------- 455
                ******************************::: .

PAX6.1          ASQYWSRIQ 489
PAX6.2          --------- 
```

Appendix 4. Sequence comparison of the protein encoded by the canonical *LsPax-6* mRNA splice variant with the protein encoded by the *LsPax-6* alternative splice

The protein sequence deduced from the canonical *LsPax-6* mRNA is referred here as Pax6.1, while the deduced protein sequence from the *LsPax-6* splice variant is referred here as Pax6.2. The PD is highlighted in yellow, while the HD is in blue. Notice that the PST domain, which is highlighted in violet, is present only in the canonical LsPax-6.

DNA sequence

```
        |   10      |   20      |   30      |   40      |   50      |   60
      D           D'
    taacccccac gcagcagcaa gcatgcagag  tcacatgacg catcaagcaa atggcgcttc 60
 61 aaccggtatg atggcctggg tacttattcg  atctttccgc ttacctctaa tacacctcca 120
121 gctcttttgg acaaatccta aatatcttgg  attcggttgt catgtactcc tcgatattac 180
181 ggaataattg ctccctacga tctgtccttg  agatctgttt cttctaatat caacaataaa 240
241 taacaataaa tgaaactctc tggtcttact  atcttcagag taaaactaaa ctcttttcta 300
301 ttttgttgga agttgctaaa cacaagtgtt  gcatgtagat ttgtctacct atagttttgt 360
361 atcttgacga tatcaaatat ttgtaaacaa  tggtctttta gagagtactt tcattttgct 420
421 cctgttggta tcaaaacata ttcttacggt  ttcattcttt attcttccag gcttaatatc 480
481 gccgggcgtc tccgtaccag tac                                        503
                                                          A
        |   10      |   20      |   30      |   40      |   50      |   60
```

Splice site predictions for 1 sequence with donor score cutoff 0.70, acceptor score cutoff 0.70 (exon/intron boundary shown in larger font):

Donor site predictions

Start	End	Score	Exon	Intron
59	73	0.89	tcaaccg	gtatgatg

Acceptor site predictions

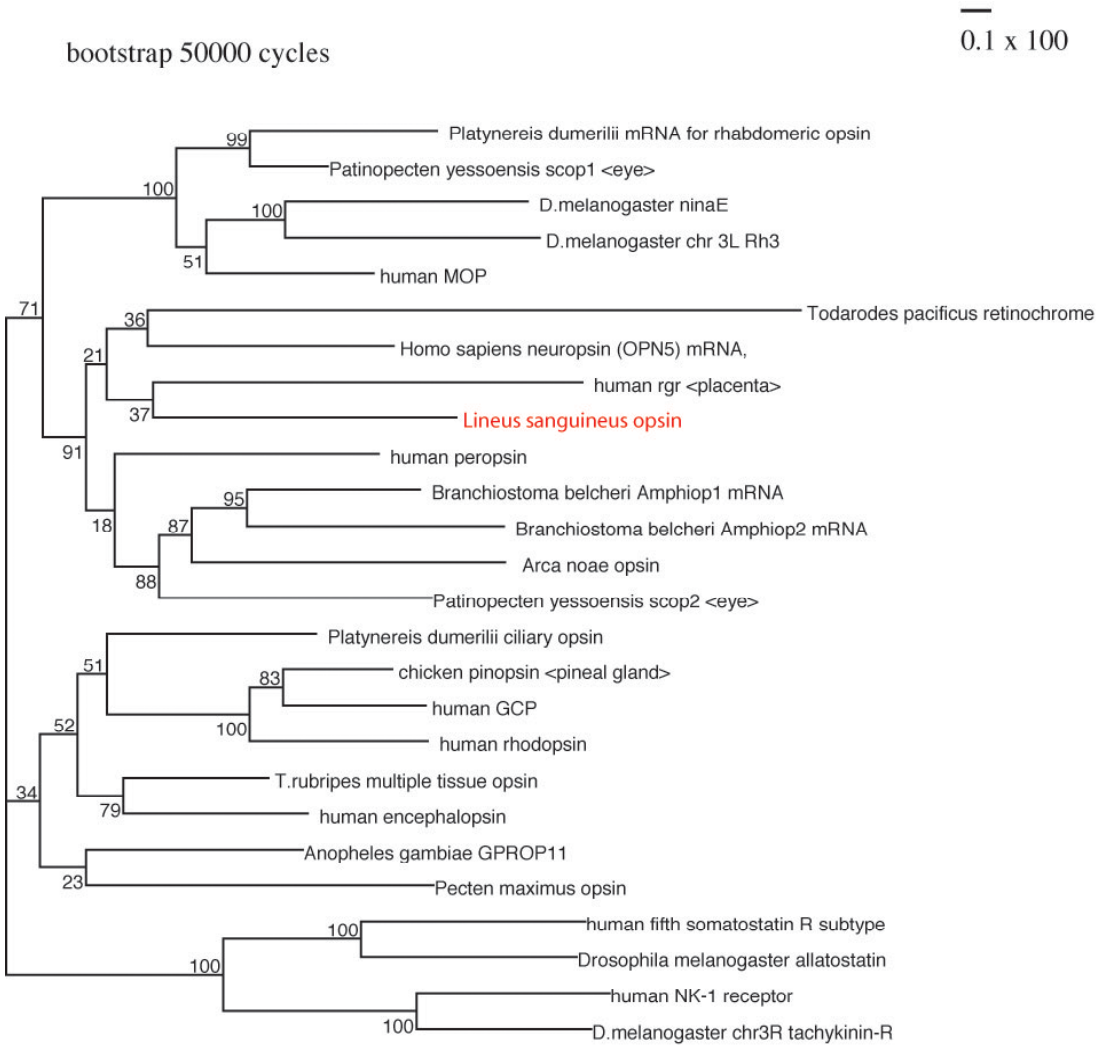
Start	End	Score	Intron	Exon
450	490	0.99	tttcattctttattcttcc	aggcttaatatcgccgggcgtc

Appendix 5. Fragment of the genomic sequence of *LsPax-6*

Exons are highlited in grey.

D: donor site

D': alternative donor site. Note that D and D' are separated by only 14 nucleotides



Appendix 6. Phylogenetic tree showing the *Ls-opsin* in relation to other invertebrate and vertebrate-type opsins.
The phylogenetic tree was constructed by using maximum likelihood method. All the amino acids from the 7TM were used.

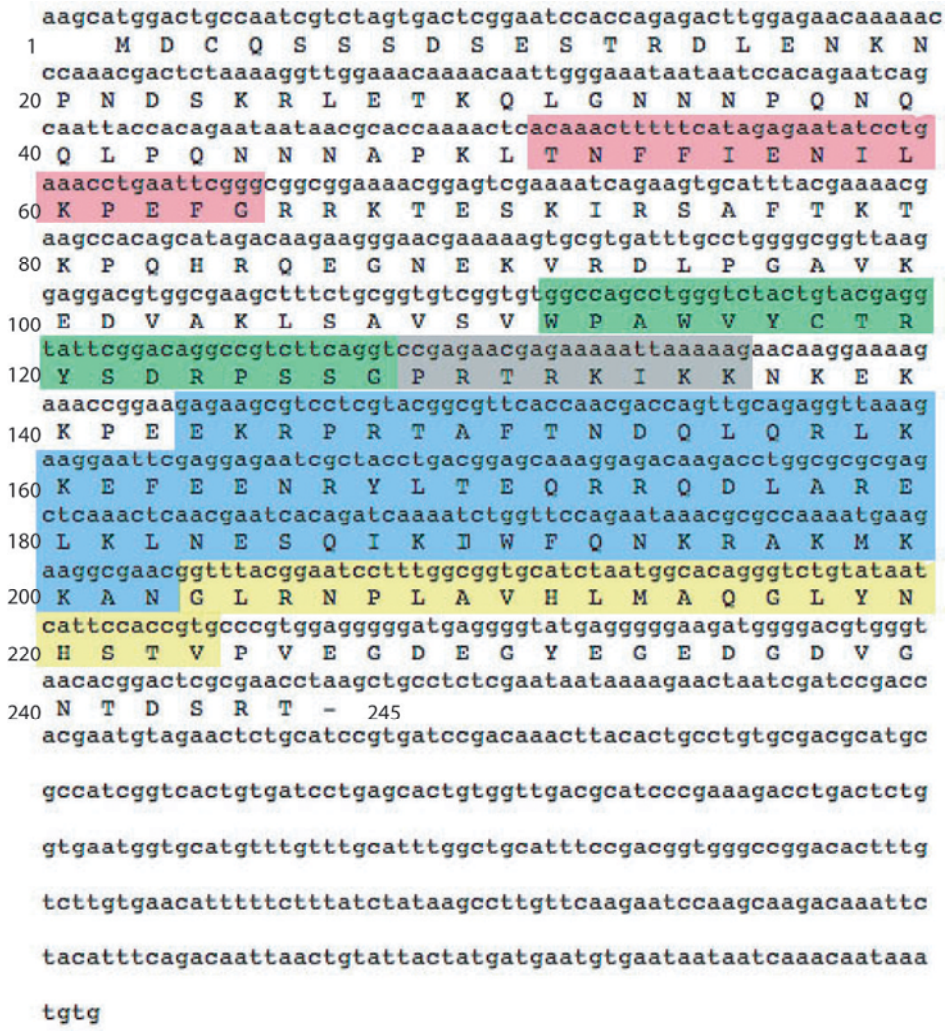
Patinopecten scop1: Gq coupled rhodopsin
Patinopecten scop2: Go coupled rhodopsin
human rgr: human Ral GDS related protein
human MOP: human melanopsin
human GCP: human green cone photoreceptor pigment
Anopheles GPROP11: "highly divergent" member of the opsin receptor family, identified by similarity to sequences in INSD and/or UniProtKB databases

```
Ls six12 / Lv six12

Score =  109 bits (272), Expect = 2e-22
Identities = 49/49 (100%), Positives = 49/49 (100%)

Query: 137 GEETSYCFKEKSRTVLKEWYAHNPYPSPREKRELAEGTGLTTMQVSNWF 185
           GEETSYCFKEKSRTVLKEWYAHNPYPSPREKRELAEGTGLTTMQVSNWF
Sbjct: 1   GEETSYCFKEKSRTVLKEWYAHNPYPSPREKRELAEGTGLTTMQVSNWF 49



Score =  220 bits (114), Expect = 1e-53
Identities = 132/143 (92%)
Strand = Plus / Plus

Query: 467 gaggagacgagttattgtttcaaggagaagtctcggacggttttgaaggaatggtaygca
           |||||||| ||||| ||||| |||||||||||||||||||||||||||||||||||| |||
Sbjct: 4   gaggagacaagttaytgttttaaggagaagtctcggacggttttgaaggaatggtacgca 63

Query: 527 cataatccatacccatctccaagggagaagagagaattagcggaaggaactggactcact
           || ||||| || || || || |||||||||||||||||||| ||||||||||||||||||
Sbjct: 64  cacaatccctatccttcgccgagggagaagagagaattagccgaaggaactggactcact 123

Query: 587 accatgcaagttagcaattggtt 609
           |||||||||||||||||||||||
Sbjct: 124 accatgcaagttagcaattggtt 146

Ls six36 / Lv six36


Score =  102 bits (254), Expect = 2e-20
Identities = 46/46 (100%), Positives = 46/46 (100%)

Query: 130 GEQKTHCFKERTRGLLREWYLQDPYPNPTKKRELAQATGLTPTQVG 175
           GEQKTHCFKERTRGLLREWYLQDPYPNPTKKRELAQATGLTPTQVG
Sbjct: 1   GEQKTHCFKERTRGLLREWYLQDPYPNPTKKRELAQATGLTPTQVG 46


Score =  235 bits (122), Expect = 3e-58
Identities = 134/140 (95%)
Strand = Plus / Plus

Query: 463 ggggaacagaaaacacactgttttaaagaacgaacacggggtttactacgggaatggtac
           ||||| ||||| || ||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 1   ggggagcagaagacgcactgttttaaagaacgaacacggggtttactacgggaatggtac 60

Query: 523 ctacaggacccttatccaaaccctacaaagaaaagagaactggcgcaggccacgggactc
           || |||||||||| ||||||||||||||||||||||||||||||||||||| ||||||
Sbjct: 61  ctgcaggacccttacccaaaccctacaaagaaaagagaactggcgcaggccacaggactc 120

Query: 583 acaccaacacaagttggaaa 602
           ||||||||||||||||||||
Sbjct: 121 acaccaacacaagttggaaa 140
```

Appendix 7. Alignment of the *Lv-Six1/2* and *Lv-Six3/6* fragments obtained by degenerated PCR and the corresponding homeobox sequences of the *Ls-Six1/2* and *Ls-Six3/6* genes

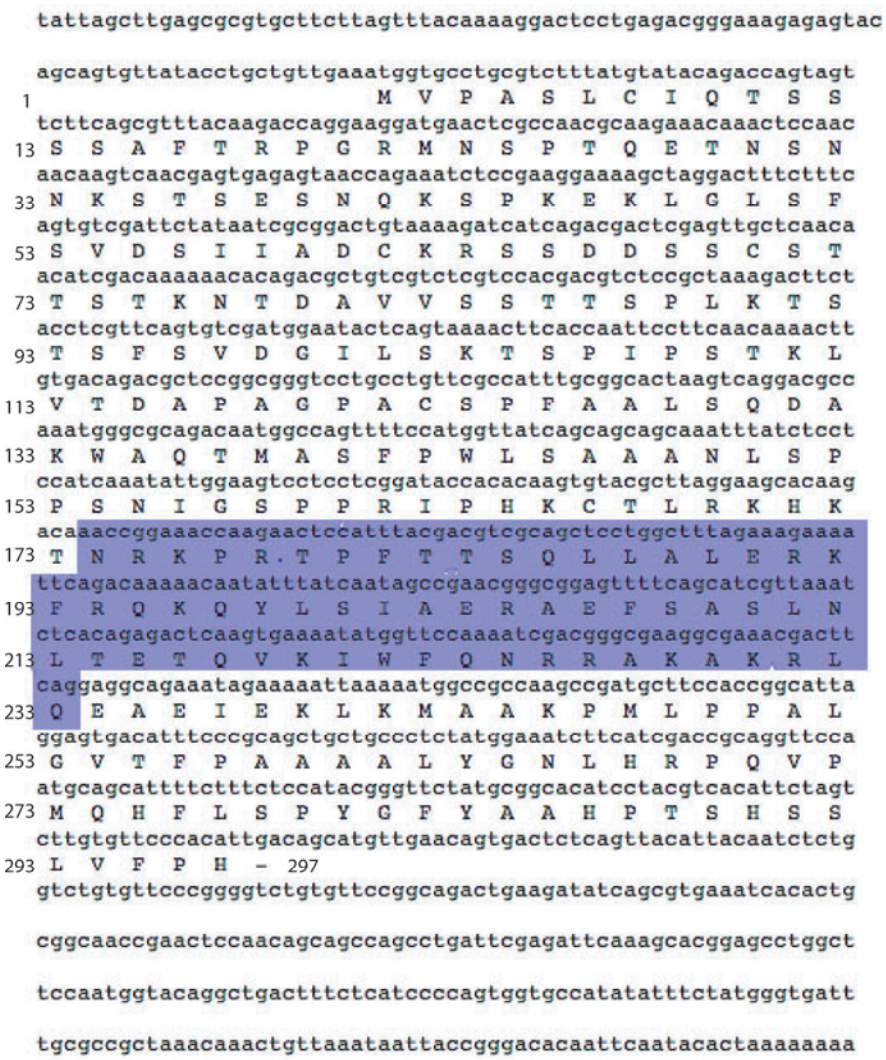
Appendix 8. The *Ls-Bmp2/4* gene and the deduced amino acid sequence of its putative ORF

The *Ls-Bmp2/4* gene encodes a putative protein of 311 aa. TheTGF β domain is highlighted in red. Notice that the ORF is probably incomplete at its 5' end.



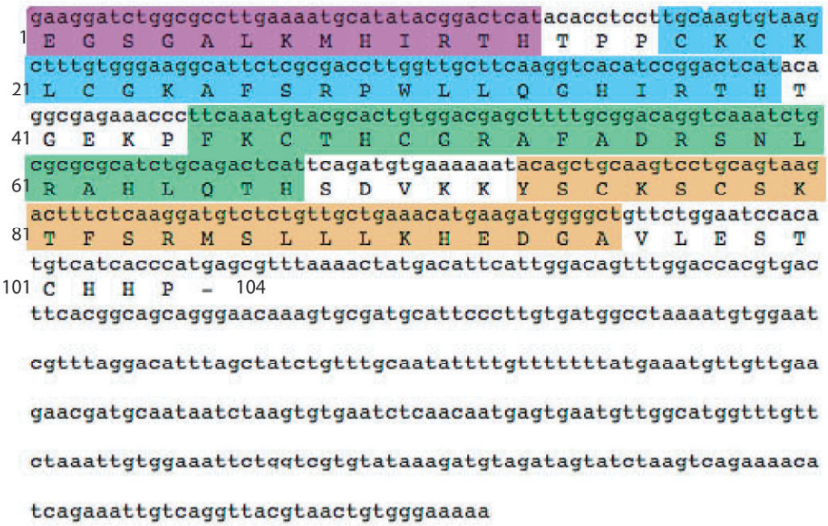
Appendix 9. The *Ls-Engrailed* gene and the deduced amino acid sequence of its putative ORF

The *Ls-Engrailed* gene encodes a putative protein of 245 aa. The Engrailed Homology (EH) 1 motif is highlighted in pink, the EH2 is in green, the EH3 is in grey. The EH4, which corresponds to the homeodomain is in blue and the EH5 is in yellow.



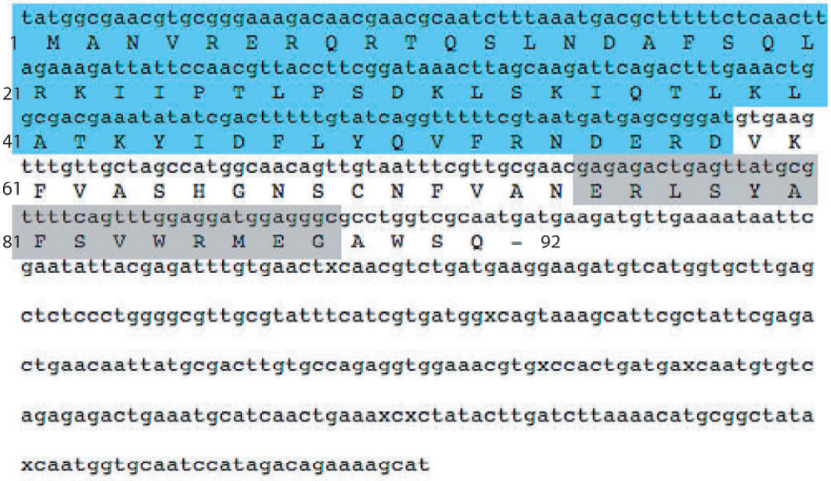
Appendix 10. The *Ls-Msx* gene and the deduced amino acid sequence of its putative ORF

The *Ls-Msx* encodes a putative protein of 297 aa. The homeodomain is highlighted in blue.



Appendix 11. Partial *Ls-Snail* gene and the deduced amino acid sequence of its putative ORF

The partial *Ls-Snail* gene encodes a putative protein of 104 aa. The zinc fingers are highlighted in color: the partial zinc finger II is in violet, the zinc finger III is in blue, the zinc finger IV is in green and the zinc finger V is in orange.



Appendix 12. Partial *Ls-Twist* gene and the deduced amino acid sequence of its putative ORF

The partial *Ls-Twist* gene encodes a putative protein of 92 aa. The incomplete bHLH domain is highlighted in blue. The so-called WR motif is highlighted in grey.

Except where otherwise indicated, this thesis is my own original work.

Véronique Charpignon

October 10th, 2006

CURRICULUM VITAE

Personal data:

Name: Véronique Charpignon
Date of birth: January 10th, 1979
Place of birth: Reims, France
Nationality: French
Martial status: Single
Privat address: 8, rue de l'étoile 68330 Huningue, France

Education:

1996	Baccalauréat
1996-1997	First year of medical studies, University of Reims
1997-1999	Diplôme d'Etudes Universitaires Générales (first two years of university degree) in Life Sciences**,** University of Reims, France Graduation date: june 1999
1999-2000	Bachelor of Sciences, Cellular biology and Physiology University of Reims, France Graduation date: june 2000
2000-2001	Master's Cellular Biology and Physiology University of Reims, France Graduation date: june 2001
2001-2002	DEA in Molecular and Cellular biology of development, Paris VI / Basel Graduation date: june 2002
2002-2006	Co-tutelle Ph.D, supervised by Prof. W.J. Gehring, University of Basel and Dr. M. Tarpin, University of Reims

Communication

2003:	Annual meeting of the French Society of Developmental Biology, Nice, France Poster: "Evolution of the body plan specification: Hints from *Lineus* studies"
2004:	Euroconference on the molecular pathways leading to regeneration, San Feliu de Guixols, Spain. Poster: "Dorso-ventral body rotations in the nermertine *Lineus sanguineus* " Annual meeting of the French Society of Genetics, Paris, France. Talk: "D-V body rotations, A-P CNS morphogenesis and D-V patterning genes in the nemertean *Lineus*"